

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:29:54 ; Search time 21.04 Seconds
(without alignments)
573.392 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSGLWYFFLCIRIKVLTG.....YMFMRVNTAKSRLLDVTLL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

all number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 3000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*

1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1082	100.0	199	AAW75956	Human cell surface
2	1082	100.0	199	AAW75957	Human cell surface
3	1082	100.0	199	AAW08026	Human activated T-
4	1082	100.0	199	AAW08731	Amino acid sequenc
5	1066.5	98.6	198	AAW92213	Human Th2-specific
6	737.5	68.2	200	AAW92212	Murine Th2-specific
7	734	67.8	200	AAW08723	Amino acid sequenc
8	722.5	66.8	200	AAW75958	Mouse cell surface
9	701	64.8	200	AAW71874	Rat cell surface p
10	696	64.3	216	AAW71875	Rat cell surface p
11	146.5	13.5	225	AAW41136	Rat CD28 protein s

12	145.5	13.4	221	21	AAW32286	Feline CD28. Feli
13	145.5	13.4	221	21	AAW32279	Cat CD28 receptor.
14	139.5	12.9	220	13	AAW20805	Human CD28 antigen
15	139.5	12.9	220	13	AAW27103	Sequence encoded b
16	139.5	12.9	220	17	AAW02131	Human CD28 cDNA pr
17	139.5	12.9	220	17	AAW91433	Human CD28 antigen
18	139.5	12.9	220	18	AAW38413	CD28. Homo sapien
19	139.5	12.9	220	18	AAW80442	Human CD28 antigen
20	139.5	12.9	220	19	AAW88451	Human cell surface
21	139.5	12.9	220	21	AAW96128	Human cell surface
22	139.5	12.9	220	21	AAW44294	Human CD28 recepto
23	134	12.4	225	20	AAW41135	Mouse CD28 protein
24	126	11.6	223	20	AAW41137	Human CD28 protein
25	120.5	11.1	367	18	AAW35862	Human CD28:IgG2a c
26	119.5	11.0	134	18	AAW35846	Human CD28 for use
27	119.5	11.0	152	16	AAW67706	CD28 extracellular
28	89.5	8.3	117	20	AAW24469	Human CD28 gene pr
29	89	8.2	321	22	AAW46999	Human OLFXY protei
30	87	8.0	330	8	AAW70445	Sequence of mouse
31	86	7.9	223	18	AAW25111	Soluble human CTLA
32	85.5	7.9	187	19	AAW29728	Soluble CTLA4 muta
33	84.5	7.8	209	21	AAW94998	Human secreted pro
34	84	7.8	223	21	AAW15129	Human CTLA-4 prote
35	81.5	7.5	187	17	AAW86945	Human T cell speci
36	81.5	7.5	187	20	AAW41130	CTLA4 receptor wit
37	81.5	7.5	187	20	AAW37615	Human CTLA recepto
38	81.5	7.5	187	20	AAW97610	Human CTLA recepto
39	81.5	7.5	211	20	AAW87560	Human CTLA4 recept
40	81.5	7.5	212	14	AAW31040	Human CTLA4 recept
41	81.5	7.5	212	15	AAW60134	CTLA4 receptor fus
42	81.5	7.5	212	16	AAW77642	Full length CTLA4
43	81.5	7.5	212	20	AAW43479	Amino acid sequenc
44	81.5	7.5	212	20	AAW81584	Human CTLA4 recept
45	81	7.5	229	11	AAW04495	HIV fusion protein

ALIGNMENTS

RESULT 1
AAW75956
ID AAW75956 standard; Protein; 199 AA.
XX
AC AAW75956;
DT 11-DEC-1998 (first entry)
XX
DE Human cell surface protein #1.
XX
KW Human; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.
XX
OS Homo sapiens.
PN WO9838216-A1.
XX
PD 03-SEP-1998.
XX
PF 27-FEB-1998; 98WO-JP00837.
XX
PR 26-FEB-1998; 98JP-0062217.
XX
PR 27-FEB-1997; 97JP-0062290.
XX
(NISR) JAPAN TOBACCO INC.
XX
PI Tamatani T, Tezuka K;
XX
DR WPI; 1998-481144/41.
XX
DR N-PSDB; AAW53198.
XX
PT Cell surface molecule expressed in thymocytes and lymphocytes and
PT mediating signal transmission and cell adhesion, and antibodies to

PT it useful in treatment of auto:immune and allergic disorders.

XX Claim 2; Page 99-101; 149pp; Japanese.

XX The present sequence represents a human cell surface protein which is expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell surface protein induces adhesion of mitogen-stimulated lymphocytes to antibodies recognising the cell surface protein. These antibodies also produce an increase in peripheral blood lymphocytes in the presence of an antibody recognising CD3 antigen. The cell surface protein contains the amino acid sequence FPPPPF in its extracellular region and the sequence YMFPM in its intracellular region. The cell surface protein can be used in the prevention and treatment of autoimmune and allergic diseases, and in the diagnosis and investigation of such disorders.

XX Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 19; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSGLWFFFLCLRIKIVLTGEINGSANYEMFIFHNGGVQILCKYPDIVQQFKMQLKGGQ 60
Db 1 mksglwffflclrikivltgeingsanyemfifhnggvqilckypdivqqfkmqlkggq 60
Qy 61 ILCDLTKTGSGNVTISKLFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
Db 61 ilcdltktgsgnvtisklfchsolnnsvsfflynldhshanyfncnlsifdpppfk 120
Qy 121 VTLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vtltggylhiesqlccqlkfwlpigcaafvvcilgclicwltkkkysssvhdnpngey 180
Qy 181 MFMRAVNTAKKSRITDVTL 199
Db 181 mfmravntakkstritdvtl 199

RESULT 2

AAW75957
ID AAW75957 standard; Protein: 199 AA.

XX AAW75957;

DT 11-DEC-1998 (first entry)

XX Human cell surface protein #2.

DE Human; cell surface protein: thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.

XX Homo sapiens.

XX WO9838216-A1.

XX 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00837.

XX 26-FEB-1998; 98JP-0062217.

XX 27-FEB-1997; 97JP-0062290.

XX (NISR) JAPAN TOBACCO INC.

XX Tamatani T, Tezuka K;

XX WPI; 1998-481144/41.

XX N-PSDB; AAV53199.

XX Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to

PT it useful in treatment of auto:immune and allergic disorders.

XX Claim 9; Page 101-105; 149pp; Japanese.

XX The present sequence represents a human cell surface protein which is expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell surface protein induces adhesion of mitogen-stimulated lymphocytes to antibodies recognising the cell surface protein. These antibodies also produce an increase in peripheral blood lymphocytes in the presence of an antibody recognising CD3 antigen. The cell surface protein contains the amino acid sequence FPPPPF in its extracellular region and the sequence YMFPM in its intracellular region. The cell surface protein can be used in the prevention and treatment of autoimmune and allergic diseases, and in the diagnosis and investigation of such disorders.

XX Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 19; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSGLWFFFLCLRIKIVLTGEINGSANYEMFIFHNGGVQILCKYPDIVQQFKMQLKGGQ 60
Db 1 mksglwffflclrikivltgeingsanyemfifhnggvqilckypdivqqfkmqlkggq 60
Qy 61 ILCDLTKTGSGNVTISKLFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
Db 61 ilcdltktgsgnvtisklfchsolnnsvsfflynldhshanyfncnlsifdpppfk 120
Qy 121 VTLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vtltggylhiesqlccqlkfwlpigcaafvvcilgclicwltkkkysssvhdnpngey 180
Qy 181 MFMRAVNTAKKSRITDVTL 199
Db 181 mfmravntakkstritdvtl 199

RESULT 3

AAW08026
ID AAW08026 standard; Protein: 199 AA.

XX AAW08026;

DT 08-JUL-1999 (first entry)

XX Human activated T-lymphocyte protein 8F4.

DE T-lymphocyte; human; 8F4; T cell co-stimulation; activated; CD4+; CD8+;
KW anticancer; antiviral; anti-asthma; immunomodulator; proliferation;
KW T cell activation; cytokine synthesis; regulatory element; B cell;
KW T cell-dependent antibody production; treatment; prevention; cancer;
KW autoimmune disease; transplant rejection; immune system regulation;
KW disorder; acquired immune deficiency syndrome; AIDS; asthma.

XX Homo sapiens.

XX WO9915553-A2.

XX 01-APR-1999.

XX 23-SEP-1998; 98WO-DE02896.

XX 11-MAY-1998; 98DE-1021060.

XX 23-SEP-1997; 97DE-1041929.

XX (DEKO-) DEUT KOCH INST ROBERT.

XX (BUND) BUNDESREPUBLIK DEUT PAUL-EHRlich-INST.

XX Krocze R;

XX WPI; 1999-276975/23.

DR N-PSDB; AAX37661.
XX Polypeptide 8F4 co-stimulates T cells and is present only on
PT activated cells
XX
XX Claim 2; Page 24; 47pp; German.
XX
CC This invention describes a novel human protein, 8F4, and its encoding
CC nucleic acid which co-stimulates T cells and is present on activated CD4+
CC CD8+ T cells but not on resting or activated B cells, granulocytes,
CC monocytes, natural killer or dendritic cells. 8F4 has anticancer,
CC antiviral, anti-asthma and immunomodulatory activity. 8F4 provides a
CC strong co-stimulatory signal for T cell activation, i.e. it amplifies
CC proliferation of T cells, synthesis of certain cytokines and other
CC regulatory agents, and improves T cell-dependent antibody production
CC by B cells. Agents that inhibit 8F4 are used to treat or prevent
CC autoimmune diseases, to prevent transplant rejection and to treat
CC disorders of immune system regulation. 8F4, or cells that express it,
CC is/are used to treat or prevent cancers, acquired immune deficiency
CC syndrome, asthma and chronic infectious diseases (e.g. hepatitis B or C).
SU Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 20; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSGLWYFFLCRLIKVLTGEINGSANYEMFTFHNGGVQILCKYPDIVQOQFMQLLKGQ 60
Db 1 mksglwyfflclrlkvlgtgeingsanyemfifhnggvqilckypdivqgfmqlkggq 60

Qy 61 ILCDLTKTGSGNTVSIKSLKFCCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
Db 61 ilcdltktkgsgntvsikslkfchsqslsnnsvsfflynlhdshanyfncnlsifdpppfk 120

Qy 121 VLTGGYLHIYESQLCCOLKFWLPIGCAAFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vltggylhiyesqlccqlkfwlpigcaafvvcilgcllcwltkkkysssvhdnpngey 180

Qy 181 MFMRVNTAKSRSLTDVTL 199
Db 181 mfmravntakksrltdvltl 199

RESULT 4
AAB08731 standard; Protein; 199 AA.
AAB08731;

XX 02-JAN-2001 (first entry)
XX Amino acid sequence of a human CRP1 polypeptide.
XX
XX CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
KW T-lymphocyte activation; type I transmembrane protein; T cell activation;
KW T cell proliferation; T-cell mediated disorder.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /note= "signal peptide"
FT Protein 21..199
FT /note= "mature protein"
FT Domain 21..140
FT /note= "extracellular domain"
FT Domain 141..161
FT /note= "predicted transmembrane domain"
FT Domain 162..199
FT /note= "intracellular domain"
XX

PN WO200046240-A2.
XX
XX 10-AUG-2000.
XX
XX 27-JAN-2000; 2000WO-US01871.
XX
XX 03-FEB-1999; 99US-0244448.
PR 08-MAR-1999; 99US-0264527.
XX (AMGE-) AMGEN INC.
XX Yoshinaga SK;
XX
XX WPI; 2000-543476/49.
DR N-PSDB; AAA64558.
XX
XX Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
XX in the treatment, prevention and diagnosis of T cell mediated disorders
XX
XX Disclosure; Fig 13A; 174pp; English.
XX
XX The present sequence represents a CRP1 (CD28 related protein-1)
XX polypeptide. The specification also describes a B7RP1 (B7 related
XX protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
XX activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are
XX predicted to be a type I transmembrane protein. The nucleic acids are
XX useful for regulating T cell activation or proliferation in an animal.
XX The polypeptides are useful for treating, preventing ameliorating or
XX diagnosing a T-cell mediated disorder in an animal. They can also be
XX used to identify test molecules that bind to the polypeptides.
XX
XX Sequence 199 AA;

Query Match 100.0%; Score 1082; DB 21; Length 199;
Best Local Similarity 100.0%; Pred. No. 3e-116;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSGLWYFFLCRLIKVLTGEINGSANYEMFTFHNGGVQILCKYPDIVQOQFMQLLKGQ 60
Db 1 mksglwyfflclrlkvlgtgeingsanyemfifhnggvqilckypdivqgfmqlkggq 60

Qy 61 ILCDLTKTGSGNTVSIKSLKFCCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
Db 61 ilcdltktkgsgntvsikslkfchsqslsnnsvsfflynlhdshanyfncnlsifdpppfk 120

Qy 121 VLTGGYLHIYESQLCCOLKFWLPIGCAAFVVCILGCLICWLTKKKYSSSVHDPNGEY 180
Db 121 vltggylhiyesqlccqlkfwlpigcaafvvcilgcllcwltkkkysssvhdnpngey 180

Qy 181 MFMRVNTAKSRSLTDVTL 199
Db 181 mfmravntakksrltdvltl 199

RESULT 5
AAY92213
ID AAY92213 standard; Protein; 198 AA.
XX
XX AAY92213;
XX
XX 10-AUG-2000 (first entry)
XX Human Th2-specific polypeptide, hl228.
XX
XX hl228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;
KW CD28; antidiabetic; immunosuppressive; neuroprotective; dermatological;
KW antipsoriatic; antiasthmatic; antiallergic; anti-viral; ophthalmological;
KW CLTA-4; nephrotropic; anti-HIV; antibacterial.
XX
XX Homo sapiens.
OS

QY 1 M K S G L W Y F F L C R I K V L T G E I N G S A N Y E M F T F H N G V Q I L C K Y P D I V Q O P K M O L L K G Q 60
 Db 1 m k p y c h v f c f l i r l i t g e i n g s a d h r m f s f n g g v q i s c k y p e t v q l k m r l f r e r e 60
 QY 61 I L C D L T K T G S G N T V S I K S L K F C H S Q L S N N S V S F F L Y N L D H S H A N Y F C N L S I F D P P P F K 120
 Db 61 v l c e i t k t g s g n a v s i k n p m l c l y h l s n n s v f f l n n p d s g g s y y f c s l s i f d p p p f q 120
 QY 121 V - T L T G G V L H I Y E S Q L C Q L K F W L P I G C A A F V V V C I L C I C W L T K K Y S S V H D P N G E 179
 Db 121 e r n l s g g y l h i y e s q l c c q l k l w p v g c a a f v v v l l f g c i l l i w f s k k y g s s v h d p n s e 180
 QY 180 Y M E M R A V N T A K K S R L T D V T 198
 Db 181 y m f m a a v n t n k k s r l a g v t 199

RESULT 7
 AAB08723
 AAB08723 standard; Protein; 200 AA.
 AAB08723;
 02-JAN-2001 (first entry)
 Amino acid sequence of a murine CRP1 polypeptide.
 CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
 T-lymphocyte activation; type I transmembrane protein; T cell activation;
 T cell proliferation; T-cell mediated disorder.
 Mus sp.
 Key Location/Qualifiers
 FH Peptide 1..20
 FT Peptide /note= "signal peptide"
 FT Protein 21..200
 FT Protein /note= "mature protein"
 FT Domain 21..145
 FT Domain /note= "extracellular domain"
 FT Domain 146..163
 FT Domain /note= "predicted transmembrane domain"
 FT Domain 164..200
 FT Domain /note= "intracellular domain"
 XX W0200046240-A2.
 XX 10-AUG-2000.
 XX 27-JAN-2000; 2000WO-US01871.
 XX 03-FEB-1999; 99US-0244448.
 PR 08-MAR-1999; 99US-0264527.
 XX (AMGE-) AMGEN INC.
 XX Yoshinaga SK;
 PI WPI; 2000-543476/49.
 DR N-PSDB; AAA64554.
 XX Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
 PT in the treatment, prevention and diagnosis of T cell mediated disorders
 PT -
 XX Claim 11; Fig 1A; 174pp; English.
 PS The present sequence represents a CRP1 (CD28 related protein-1)
 CC polypeptide. The specification also describes a B7RP1 (B7 related
 CC protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
 CC activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are
 CC predicted to be a type I transmembrane protein. The nucleic acids are
 CC useful for regulating T cell activation or proliferation in an animal.

CC The polypeptides are useful for treating, preventing ameliorating or
 CC diagnosing a T-cell mediated disorder in an animal. They can also be
 CC used to identify test molecules that bind to the polypeptides.
 XX Sequence 200 AA;
 SQ
 Query Match 67.8%; Score 734; DB 21; Length 200;
 Best Local Similarity 70.4%; Pred. No. 2.9e-76;
 Matches 138; Conservative 19; Mismatches 35; Indels 4; Gaps 2;
 QY 7 Y F - - - F L C L R I K V L T G E I N G S A N Y E M F T F H N G V Q I L C K Y P D I V Q O P K M O L L K G Q I L C 63
 Db 4 y f c r v f c f l i r l i t g e i n g s a d h r m f s f n g g v q i s c k y p e t v q l k m r l f r e r e v i c 63
 QY 64 D L T K T G S G N T V S I K S L K F C H S Q L S N N S V S F F L Y N L D H S H A N Y F C N L S I F D P P P F K V - T 122
 Db 64 e l t k t g s g n a v s i k n p m l c l y h l s n n s v f f l n n p d s g g s y y f c s l s i f d p p p f q e r n 123
 QY 123 L T G G V L H I Y E S Q L C Q L K F W L P I G C A A F V V V C I L C I C W L T K K Y S S V H D P N G E Y M F 182
 Db 124 l s g g y l h i y e s q l c c q l k l w p v g c a a f v v v l l f g c i l l i w f s k k y g s s v h d p n s e y m f 183
 QY 183 M R A V N T A K K S R L T D V T 198
 Db 184 m a a v n t n k k s r l a g v t 199

RESULT 8
 AAW75958
 ID AAW75958 standard; Protein; 200 AA.
 XX AC AAW75958;
 XX 11-DEC-1998 (first entry)
 DT Mouse cell surface protein.
 DE Mouse cell surface protein.
 XX KW Mouse; cell surface protein; thymocyte; lymphocyte; cell adhesion;
 KW signal transmission; autoimmune disorder; allergy; diagnosis;
 XX mitogen-stimulated.
 OS Mus sp.
 XX W09838216-A1.
 PN 03-SEP-1998.
 XX 27-FEB-1998; 98WO-JP00837.
 XX 26-FEB-1998; 98JP-0062217.
 PR 27-FEB-1997; 97JP-0062290.
 XX (NISB) JAPAN TOBACCO INC.
 PA Tamatani T, Tezuka K;
 PI WPI; 1998-481144/41.
 DR N-PSDB; AAV53200.
 XX Cell surface molecule expressed in thymocytes and lymphocytes and -
 PT mediating signal transmission and cell adhesion, and antibodies to
 PT it useful in treatment of auto-immune and allergic disorders.
 XX Claim 9; Page 110-112; 149pp; Japanese.
 PS The present sequence represents a mouse cell surface protein which is
 CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
 CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
 CC antibodies recognising the cell surface protein. These antibodies also
 CC produce an increase in peripheral blood lymphocytes in the presence of
 CC an antibody recognising CD3 antigen. The cell surface protein contains
 CC the amino acid sequence FDPFPF in its extracellular region and the

CC * sequence YMFH in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.

XX SQ Sequence 200 AA;

Query Match 66.8%; Score 722.5; DB 19; Length 200;
Best Local Similarity 68.3%; Pred. No. 6e-75;
Matches 136; Conservative 20; Mismatches 42; Indels 1; Gaps 1;

QY 1 MKSGLWFFFLCRLIKVLTGEINGSANYEMFIHNGGVQLCKYPDVIQOQFKMOLLKGGQ 60
DB 1 mkpyfchvfvcflirltgeingsadhrmfshfnggqvicskypetvqqlkmrlfrere 60
QY 61 ILCDLTKSGNTVTSIKSLKFCQSLSNNSVSFFLYNLDHSHANYFCNLSTFDPPPFK 120
DB 61 vlceltkgsgnavslnkpmclciyhlnsnsvsfllanpdsqgsyvfcslsfdpppfq 120
121 V-TTGGYLHIYESQLCCQLKFWLPICGAFFVVVVCILGCLICWLTKKKYSSSVHDPNGE 179
121 ernlsgylhiyesqlccqlkfwlpvqgpaifvvvllfvcililwfskkygssvndpnse 180
QY 180 YMFRAVNTAKSKRLTDVT 198
DB 181 ymfmaavntnkkksrlagvt 199

RESULT 9

AAW71874
ID AAW71874 standard; Protein; 200 AA.

XX AC AAW71874;

DT 11-DEC-1998 (first entry)

XX DE Rat cell surface protein #1.

XX KW Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.

XX OS Rattus sp.

XX PN WO9838216-A1.

XX PD 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00837.

PR 26-FEB-1998; 98JP-0062217.

PR 27-FEB-1997; 97JP-0062290.

XX PA (NISR) JAPAN TOBACCO INC.

XX PI Tamatani T, Tezuka K;

XX WPI; 1998-481144/41.

DR N-PSDB; AAV61357.

XX Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to
PT it useful in treatment of autoimmune and allergic disorders.

XX Claim 9; Page 106-109; 149pp; Japanese.

XX The present sequence represents a rat cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FPPPPF in its extracellular region and the

CC sequence YMFH in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.

XX SQ Sequence 200 AA;

Query Match 64.8%; Score 701; DB 19; Length 200;
Best Local Similarity 67.9%; Pred. No. 1.8e-72;
Matches 133; Conservative 17; Mismatches 42; Indels 4; Gaps 2;

QY 7 YF---FLFCURIKVLTGEINGSANYEMFIHNGGVQLCKYPDVIQOQFKMOLLKGGQILC 63
DB 4 yfscvfvfcflirltgeindlanhrmfshdggqvicscypetvqqlkmlfkdvrlc 63
QY 64 DLTKTKSGNTVTSIKSLKFCQSLSNNSVSFFLYNLDHSHANYFCNLSTFDPPPF-KVT 122
DB 64 dlctkkgsgntvslnkpmcspyglnsnsvafldnadsgqsyfclalsfdpppfqekn 123
QY 123 LTGGYLHIYESQLCCQLKFWLPICGAFFVVVVCILGCLICWLTKKKYSSSVHDPNGEYMF 182
DB 124 lsggylhiyesqlccqlkfwlpvqgpaifvvvllfvcililwfskkyrsvvhdnpseymf 183
QY 183 MRVAVNTAKSKRLTDVT 198
DB 184 maaavntnkkksrlagmt 199

RESULT 10

AAW71875
ID AAW71875 standard; Protein; 216 AA.

XX AC AAW71875;

DT 11-DEC-1998 (first entry)

XX DE Rat cell surface protein #2.

XX KW Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transmission; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.

XX OS Rattus sp.

XX PN WO9838216-A1.

XX PD 03-SEP-1998.

XX 27-FEB-1998; 98WO-JP00837.

PR 26-FEB-1998; 98JP-0062217.

PR 27-FEB-1997; 97JP-0062290.

XX PA (NISR) JAPAN TOBACCO INC.

XX PI Tamatani T, Tezuka K;

XX WPI; 1998-481144/41.

DR N-PSDB; AAV61358.

XX Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to
PT it useful in treatment of autoimmune and allergic disorders.

XX Claim 9; Page 112-115; 149pp; Japanese.

XX The present sequence represents a rat cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FPPPPF in its extracellular region and the

FT Modified-site /label= N-linked_glycosylation
 FT 71...73
 FT Modified-site /label= N-linked_glycosylation
 FT 92...94
 FT Modified-site /label= N-linked_glycosylation
 FT 105...107
 FT Modified-site /label= N-linked_glycosylation
 FT 129...131
 FT Modified-site /label= N-linked_glycosylation
 FT 153...179
 FT Region /label= transmembrane
 XX
 XX
 PN W09201049-A.
 XX
 PD 23-JAN-1992.
 XX
 XX 15-JUL-1990; 90WO-US04986.
 PF
 XX 13-JUL-1990; 90US-0553759.
 PR
 XX (GEHO-) GEN HOSPITAL CORP.
 PI Seed B, Aruffo A, Amiot M;
 XX
 XX WPI; 1992-056864/07.
 DR N-PSDB; AAQ21167.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 immunotherapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 XX Example 3; Fig 7; 160pp; English.
 PS
 XX The CD28 antigen amino acid sequence was predicted from the
 CC nucleotide sequence of a cDNA clone isolated from a human
 CC lymphoblastoid cell line JY library using the antibody enrichment
 CC method (see AAQ21167). The sequence of the CD28 antigen has
 CC substantial homology with mouse and rabbit immunoglobulin
 CC heavy-chain variable regions over a domain spanning almost the
 CC entire extracellular portion of CD28.
 XX
 XX Sequence 220 AA;
 SQ

Query Match 12.9%; Score 139.5; DB 13; Length 220;
 Best Local Similarity 26.0%; Pred. No. 6.2e-08;
 Matches 45; Conservative 30; Mismatches 73; Indels 25; Gaps 7;
 QV 30 MFIFNGVQILCKYPD--IVQOFKMLKGGQILCDLTKTSGSN---TVSIKSLKFCH 84
 28 mlvaydnvnlsckysynlfrfrefraslhkgldsavevcvvygnyssqqlqvysktgfncd 87
 QY 85 SOLSNSVSFFLYNLDHSHANYFCNLSIFDPPPP--KVTLTGGYLHYVESOLCCQLKFW 142
 Db 88 gklgnesvtfylqnlvvnqtdiyfckievmpypdyldneksngtlihvkgkhlcpplf- 146
 QY 143 LPIGCAAFVVCILCILIC-----WLTKKYSVSSVHDHPNGEYMPM 183
 Db 147 -pgpskptfwlvvvgglacysllvtvafilfwvrskr-srllh---sdymnm 194
 RESULT 15
 ID AAR27103
 XX AAR27103 standard; Protein; 220 AA.
 AC
 XX AAR27103;
 XX
 DT 25-FEB-1993 (first entry)
 XX
 DE Sequence encoded by the CD28 gene.
 XX
 XX CD28 protein; T-cell immune response; Tp44; differentiation antigen;
 KW membrane protein.
 XX

OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /label= signal
 FT 153..179
 FT Region /label= Transmembrane region
 FT
 XX
 PN W09215671-A.
 XX
 XX 17-SEP-1992.
 XX
 XX 09-MAR-1992; 92WO-US01867.
 PF
 XX 08-MAR-1991; 91US-0666809.
 PR
 XX (CVTO-) CYTOMED INC.
 PA
 XX KO JL, IP SH;
 PI
 XX WPI; 1992-331717/40.
 DR N-PSDB; AAQ28837.
 XX
 PT Soluble proteins binding to B7 proteins and block antigen presenting
 FT cells - are useful in treating T-cell mediated immunosuppression
 FT diseases e.g. transplant rejection, autoimmune diseases etc.
 XX
 PS Disclosure; Fig 1; 43pp; English.
 XX
 CC The cDNA sequence in AAQ28837 is derived from Aruffo A and Seed B,
 CC PNAS USA 84, 8573-77, 1987. The CD28 cDNA encodes a protein of 220
 CC AAs (AAR27103). It contains a signal peptide from AAs 1-18 which is
 CC cleaved during the maturation of CD28 protein. The transmembrane
 CC region is AAs 153-79. For the production of soluble CD28 protein,
 CC the transmembrane region is deleted. The CD28 protein derived from
 CC p9-MIT comprises residues 19-157 and is not soluble; the CD28
 CC protein derived from p9-ClT comprises residues 19-151 and is soluble.
 CC MIT and ClT are primers. T7 primer is used as the 5' primer to pair
 CC with any 3' primer, eg, MIT and ClT, in a PCR reaction for the
 CC procurement of truncated CD28 gene which would direct the production
 CC of soluble CD28 protein. A compound containing residues 19-151 of
 CC CD28 protein is claimed.
 XX
 XX Sequence 220 AA;
 SQ
 Query Match 12.9%; Score 139.5; DB 13; Length 220;
 Best Local Similarity 26.0%; Pred. No. 6.2e-08;
 Matches 45; Conservative 30; Mismatches 73; Indels 25; Gaps 7;
 QY 30 MFIFNGVQILCKYPD--IVQOFKMLKGGQILCDLTKTSGSN---TVSIKSLKFCH 84
 Db 28 mlvaydnvnlsckysynlfrfrefraslhkgldsavevcvvygnyssqqlqvysktgfncd 87
 QY 85 SOLSNSVSFFLYNLDHSHANYFCNLSIFDPPPP--KVTLTGGYLHYVESOLCCQLKFW 142
 Db 88 gklgnesvtfylqnlvvnqtdiyfckievmpypdyldneksngtlihvkgkhlcpplf- 146
 QY 143 LPIGCAAFVVCILCILIC-----WLTKKYSVSSVHDHPNGEYMPM 183
 Db 147 -pgpskptfwlvvvgglacysllvtvafilfwvrskr-srllh---sdymnm 194
 Search completed: October 30, 2001, 20:39:43
 Job time: 589 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:34:54 ; Search time 13.57 Seconds
(without alignments)
301.951 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSGLWFFFLCLRIKVLG.....YMFMRVNTAKSKRLTDVTL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

1 number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163	15.1	218	3	US-08-228-208A-20
2	146.5	13.5	225	1	US-08-505-058-4
3	146.5	13.5	225	2	US-08-459-818-24
4	146.5	13.5	225	2	US-08-889-666-24
5	146.5	13.5	225	2	US-08-465-078-24
6	146.5	13.5	225	2	US-08-725-776-24
7	146.5	13.5	225	2	US-08-488-062-24
8	140	12.9	218	3	US-08-228-208A-19
9	134.5	12.4	220	3	US-08-228-208A-21
10	134	12.4	225	1	US-08-505-058-3
11	134	12.4	225	2	US-08-459-818-23
12	134	12.4	225	2	US-08-889-666-23
13	134	12.4	225	2	US-08-465-078-23
14	134	12.4	225	2	US-08-725-776-23
15	134	12.4	225	2	US-08-488-062-23
16	126	11.6	223	1	US-08-505-058-5
17	126	11.6	223	2	US-08-459-818-25
18	126	11.6	223	2	US-08-889-666-25
19	126	11.6	223	2	US-08-465-078-25
20	126	11.6	223	2	US-08-725-776-25
21	126	11.6	223	2	US-08-488-062-25
22	120.5	11.1	367	3	US-08-630-172-19
23	119.5	11.0	134	3	US-08-630-172-3
24	93	8.6	221	3	US-08-228-208A-22
25	89.5	8.3	117	2	US-08-529-878B-39
26	87	8.0	330	2	US-08-332-562A-81
27	87	8.0	330	2	US-08-332-562A-134

28 84 7.8 223 3 US-08-228-208A-17 Sequence 17, Appl
29 84 7.8 283 2 US-08-332-562A-136 Sequence 136, App
30 81.5 7.5 187 1 US-08-067-684-14 Sequence 14, Appl
31 81.5 7.5 187 1 US-08-008-898-14 Sequence 14, Appl
32 81.5 7.5 187 2 US-08-459-818-14 Sequence 14, Appl
33 81.5 7.5 187 2 US-08-889-666-14 Sequence 14, Appl
34 81.5 7.5 187 2 US-08-465-078-14 Sequence 14, Appl
35 81.5 7.5 187 2 US-08-725-776-14 Sequence 14, Appl
36 81.5 7.5 187 3 US-08-488-062-14 Sequence 14, Appl
37 81.5 7.5 187 3 US-08-228-208A-14 Sequence 14, Appl
38 81.5 7.5 187 3 PCT-US95-06726-36 Sequence 36, Appl
39 80.5 7.4 223 3 US-08-228-208A-18 Sequence 18, Appl
40 80.5 7.4 234 1 US-08-505-058-1 Sequence 1, Appl
41 80.5 7.4 234 2 US-08-459-818-21 Sequence 21, Appl
42 80.5 7.4 234 2 US-08-889-666-21 Sequence 21, Appl
43 80.5 7.4 234 2 US-08-465-078-21 Sequence 21, Appl
44 80.5 7.4 234 2 US-08-725-776-21 Sequence 21, Appl
45 80.5 7.4 234 2 US-08-488-062-21 Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-228-208A-20
; Sequence 20, Application US/08228208A
; Patent No. 6090914
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Wallace, Philip M.
; TITLE OF INVENTION: CTAA4/CD281g HYBRID FUSION
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/228,208A
; FILING DATE: 15-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/008,898
; FILING DATE: 22-JAN-1993
; APPLICATION NUMBER: 07/723,617
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-30US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310 445-1140
; TELEFAX: 310 445-9031
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-228-208A-20

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:36:04 ; Search time 13.62 Seconds
(without alignments)
1112.976 Million cell updates/sec

Title: US-09-383-551b-2
Perfect score: 1082
Sequence: 1 MKSGLWFFFLCRLIKVLGT.....YMPRAVNTAKKSLDVTTL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

1 number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1082	100.0	199	2 J78540	inducible T-cell c
2	701	64.8	200	2 JC7397	activation-inducib
3	696	64.3	216	2 JC7396	activation-inducib
4	152	14.0	218	2 S24413	T-cell surface gly
5	150.5	13.9	221	2 I46689	CD28 precursor - r
6	139.5	12.9	173	2 I46197	cell surface prote
7	139.5	12.9	220	1 RWHU28	T-cell surface gly
8	138	12.8	218	2 A43523	T-cell surface gly
9	103.5	9.6	988	2 T18986	hypothetical prote
10	97	9.0	221	2 S25168	CHT28 protein - ch
11	91.5	8.5	383	2 T21946	hypothetical prote
12	87	8.0	330	2 A40071	Fc gamma (IlgG) rec
13	87	8.0	330	2 I49660	Fc-gamma-1/gamma-2
14	87	8.0	563	2 T32479	hypothetical prote
15	86.5	8.0	223	2 A29063	cytotoxic T-lympho
16	86	7.9	223	2 T09536	cytotoxic T-lympho
17	84	7.8	283	1 FCWSGI	Fc gamma (IlgG) rec
18	84	7.8	285	2 D69440	conserved hypothet
19	83.5	7.7	186	2 S08614	cytotoxic T-lympho
20	82.5	7.6	248	1 Q0BE4L	probable glycoprot
21	81	7.5	276	2 S20690	31.6K hypothetical
22	81	7.5	580	2 JC5895	killer cell inhibi
23	81	7.5	1584	2 S57161	hypothetical prote
24	80.5	7.4	635	2 JC5896	killer cell inhibi
25	78	7.2	231	2 T23136	hypothetical prote
26	77.5	7.2	80	2 F86027	hypothetical prote
27	77.5	7.2	220	2 A48581	37K proline-rich s
28	77.5	7.2	1070	2 JC4593	protein-tyrosine k
29	77	7.1	338	2 T34364	hypothetical prote

30	77	7.1	2104	2 H86127	hypothetical prote
31	76.5	7.1	223	2 I46696	CTLA-4 precursor -
32	76.5	7.1	323	2 S06946	Fc gamma (IlgG) rec
33	76.5	7.1	1237	2 A54080	protein-tyrosine-p
34	76	7.0	301	2 I54209	hypothetical prote
35	76	7.0	317	2 JL0118	Fc gamma (IlgG) rec
36	76	7.0	546	2 S52053	cytochrome-c oxida
37	75.5	7.0	261	2 S29360	Fc gamma (IlgG) rec
38	75.5	7.0	277	2 T21330	hypothetical prote
39	75.5	7.0	2135	2 T14602	variant-specific s
40	75	6.9	1132	1 QSBPL	host specificity p
41	74.5	6.9	235	2 I50610	T-cell surface gly
42	74.5	6.9	415	2 H86204	hypothetical prote
43	74.5	6.9	445	2 T50802	serine/threonine p
44	74.5	6.9	682	2 A35969	heparin-binding gr
45	74.5	6.9	763	2 S17998	gene COX1 intron 4

ALIGNMENTS

RESULT 1

S78540
inducible T-cell co-stimulator ICOS precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 07-May-1999
C:Accession: S78540; S78748; S78749
R:Kroczek, R.
submitted to the Protein Sequence Database, June 1998
A:Reference number: S78540
A:Accession: S78540
A:Molecule type: mRNA
A:Residues: 1-199 <KRO>
A:Experimental source: cell line MOLT-4V
R:Hutloff, A.; Dittlich, A.M.; Beier, K.C.; Eijaschewitsch, B.; Kraft, R.; Anagnostop
Nature 397, 263-266, 1999
A:Title: ICOS is an inducible T-cell co-stimulator structurally and functionally rela
A:Reference number: S78748; MUID:99127892
A:Accession: S78748
A:Molecule type: mRNA
A:Residues: 1-199 <HUTI>
A:Experimental source: cell line MOLT-4V
A:Accession: S78749
A:Molecule type: protein
A:Residues: 'X',193-198 <HUT2>
A:Experimental source: cell line MOLT-4V
C:Complex: homodimer
C:Superfamily: immunoglobulin homology
C:Keywords: dimer; glycoprotein; T-cell; transmembrane protein
F;1-19/Domain: (or 1-20) signal sequence #status predicted <SIG>
F;20-199/Product: (or 21-199) inducible costimulator ICOS #status predicted <MAT>
F;21-138/Domain: extracellular #status predicted <EXT>
F;26-132/Domain: immunoglobulin homology <IMM>
F;139-164/Domain: transmembrane #status predicted <TM>
F;165-199/Domain: intracellular #status predicted <INT>
F;23,89,110/Binding site: carbohydrate (Asn) #status predicted

Query Match 100.0%; Score 1082; DB 2; Length 199;
Best Local Similarity 100.0%; Pred No. 1.2e-97;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKSGLWFFFLCRLIKVLGTGEINGSANYEMFIFHNGGVQILCKYPDIVQOFKMLLKGQ	60
Db	1	MKSGLWFFFLCRLIKVLGTGEINGSANYEMFIFHNGGVQILCKYPDIVQOFKMLLKGQ	60
QY	61	ILCDLTKTSGNGTVTSKLFCHSLSNNVSFFFLNLDHSHANYFCNLISFDPPPFK	120
Db	61	ILCDLTKTSGNGTVTSKLFCHSLSNNVSFFFLNLDHSHANYFCNLISFDPPPFK	120
QY	121	VTLTGGVLIHYESQLCCQLKFWLPICGAAFWVCILGCIILICWLTKKKYSSSVHDPNGEY	180
Db	121	VTLTGGVLIHYESQLCCQLKFWLPICGAAFWVCILGCIILICWLTKKKYSSSVHDPNGEY	180

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:40:10 ; Search time 10.55 Seconds
(without alignments)
646.146 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MKSLGWFFFLCLIKVLTG.....YMFMRVNTAKKSLRLTDVTL 199

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

1 number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	14.7	218	1	CD28_RAT
2	150.5	13.9	221	1	CD28_RABIT
3	141.1	13.0	219	1	CD28_BOVIN
4	139.5	12.9	220	1	CD28_HUMAN
5	138	12.8	218	1	CD28_MOUSE
6	97	9.0	221	1	CD28_CHICK
7	87	8.0	230	1	FCG2_MOUSE
8	86.5	8.0	223	1	CTLA_MOUSE
9	86	7.9	223	1	CTLA_HUMAN
10	82.5	7.6	248	1	VGL2_EBV
11	81	7.5	1584	1	YJ9G_YEAST
12	77.5	7.2	324	1	CRFB_SHEEP
13	77.5	7.2	1070	1	PTK7_HUMAN
14	76.5	7.1	223	1	CTLA_RABIT
15	76.5	7.1	323	1	FCG3_HUMAN
16	76	7.0	317	1	FCGA_HUMAN
17	76	7.0	448	1	BCN1_MOUSE
18	76	7.0	450	1	BCN1_HUMAN
19	76	7.0	466	1	ALAA_RABIT
20	76	7.0	562	1	ALAA_MOUSE
21	75.5	7.0	261	1	FCG3_MOUSE
22	75	6.9	462	1	CATC_MOUSE
23	75	6.9	1132	1	VHSJ_LAMBD
24	74	6.8	344	1	OM40_DROME
25	74	6.8	367	1	NULM_PODAN
26	73.5	6.8	285	1	FCG2_RAT
27	73.5	6.8	322	1	CRFB_MOUSE
28	73.5	6.8	360	1	CCPH_HSVSA
29	73.5	6.8	496	1	GRB_MOUSE
30	73.5	6.8	496	1	GRB_RAT
31	73.5	6.8	497	1	GRB_HUMAN
32	73.5	6.8	1162	1	BXEN_CLOBU
33	73	6.7	275	1	YD84_YEAST

34 73 6.7 916 1 SCRB_LIMPO Q25386 limulus pol
35 72.5 6.7 654 1 BFR2_HUMAN Q01742 homo sapien
36 72 6.7 561 1 AIAD_RAT P23944 rattus norv
37 72 6.7 569 1 YL53_YEAST Q06567 saccharomyc
38 71.5 6.6 634 1 YC33_EUGGR P31916 euglena gra
39 71.5 6.6 1051 1 PTK7_CHICK Q91048 gallus gall
40 71 6.6 212 1 KITH_ENCCU O96720 encephalito
41 71 6.6 466 1 ALAA_BOVIN P18130 bos taurus
42 70.5 6.5 524 1 VLI_HPV58 P26535 human papil
43 70 6.5 322 1 CRFB_HUMAN P24387 homo sapien
44 70 6.5 462 1 CATC_RAT P80067 rattus norv
45 70 6.5 859 1 CHS1_SCHPO P30597 schizosacch

ALIGNMENTS

RESULT 1
CD28_RAT
ID CD28_RAT STANDARD; PRT; 218 AA.
AC P31042;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE T-CELL-SPECIFIC SURFACE GLYCOPROTEIN CD28 PRECURSOR.
GN CD28.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA; TISSUE=Lymphoid;
RX MEDLINE=92104640; PubMed=1309509;
RA Clark G.J., Dallman M.J.;
RT Identification of a cDNA encoding the rat CD28 homologue.*;
RL Immunogenetics 35:54-57(1992).
CC -!- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
AND B7-2 (B70).
CC -!- SUBUNIT: HOMODIMER, LINKED BY A DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. STRONGEST
SIMILARITY TO CTLA-4.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; X55288; CAA39003.1; -.
DR PIR; S38722; S38722.
KW Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 218 T-CELL-SPECIFIC SURFACE GLYCOPROTEIN
CD28.
FT DOMAIN 20 150 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 151 177 POTENTIAL.
FT DOMAIN 178 218 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 138 IG-LIKE V-TYPE DOMAIN.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 218 AA; 25170 MW; 2E151C8F324C0B6E CRC64;

Query Match 14.7%; Score 159; DB 1; Length 218;
Best Local Similarity 26.5%; Pred. No. 1.8e-08;
Matches 41; Conservative 30; Mismatches 66; Indels 18; Gaps 7;

QY 30 MFIHNGGVOILCKYPD--IVQOFKMQLLKGGQILCDLTKTGSGNTVSISLK-----F 82
 Db 29 LLVVDNNEVSLSCRYNLLAKEFRASLYKG--VNSDVEVCVCGNGFTYQOPFRPNVNGEN 86
 QY 83 CHSOLSNNSVSFFLYNLDHSHANYFCNLISFDPPPF--KVTLTGGYLHIYESQLC---C 137
 Db 87 CDGDFNETVTFRLMNLVDNHTDIYFCKIEVMYPPPYLDNEKSNGTIIHKEKHLCHAQT 146
 QY 138 OLKFWLPIGCAAFVVC--ILGCILIC--WLTKKK 168
 Db 147 SPKLFVPLVAVGVLICGLLVTVTLCTIIWTNSRR 181

RESULT 2
 CD28_RABIT STANDARD; PRT; 221 AA.
 AC Q28071.
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 01-OCT-1996 (Rel. 34, Last annotation update)
 T-CELL-SPECIFIC SURFACE GLYCOPROTEIN CD28 PRECURSOR.
 CD28.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B/J X CHBB:HM;
 RX MEDLINE=95369849; PubMed=7642234;
 RA Isono T., Seto A.;
 RT Cloning and sequencing of the rabbit gene encoding T-cell
 RT costimulatory molecules.*;
 RL Immunogenetics 42:217-220(1995).
 CC -|- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
 CC AND B7-2 (B70) (BY SIMILARITY).
 CC -|- SUBUNIT: HOMODIMER, LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -|- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. STRONGEST
 CC SIMILARITY TO CTLA-4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC

EMBL; D49841; BAA08641.1; -;
 Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane.
 KW SIGNAL 1 19 BY SIMILARITY.
 FT CHAIN 20 221 T-CELL-SPECIFIC SURFACE GLYCOPROTEIN
 FT CD28.
 FT DOMAIN 20 150 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 151 177 POTENTIAL.
 FT DOMAIN 178 221 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 29 138 IG-LIKE V-TYPE DOMAIN.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 221 AA; 25307 MW; 3DF652C9CFC14F13 CRC64;
 Query Match 13.9%; Score 150.5; DB 1; Length 221;
 Best Local Similarity 23.8%; Pred. No. 1.2e-07;
 Matches 38; Conservative 27; Mismatches 70; Indels 25; Gaps 6;

QY 30 MFIHNGGVOILCKYPD--IVQOFKMQLLKGGQILCDLTKTGSGNTVSISL 80
 Db 29 MLVYNNNEVNLCKTYTNLFSEKFRASLYKGADSAVEVCVNGNFSHPHFHSTTGFN-- 86

QY 81 KFCHSOLSNNSVSFFLYNLDHSHANYFCNLISFDPPPF--KVTLTGGYLHIYESQLC-- 136
 Db 87 --CDGLGNETVTFYLNLYVNOTDIYFCKIEVMYPPPYLDNEKSNGTIIHVEQHFPCA 144
 QY 137 -----COLKFWLPI---GCAAFVVCILGCILICWLTKKK 168
 Db 145 HPSPKSSTLFWLVVVGAVLAFYFSLVTLVAFSCWMSKK 184

RESULT 3
 CD28_BOVIN STANDARD; PRT; 219 AA.
 AC Q28071.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 01-NOV-1997 (Rel. 35, Last annotation update)
 T-CELL-SPECIFIC SURFACE GLYCOPROTEIN CD28 PRECURSOR.
 GN CD28.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96186531; PubMed=8606060;
 RA Parsons K.R., Young J.R., Collins R.A., Howard C.J.;
 RT "cattle CTLA-4, CD28 and chicken CD28 bind CD86: MYPPPY is not
 RT conserved in cattle CD28.*";
 RL Immunogenetics 43:388-391(1996).
 CC -|- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
 CC AND B7-2 (B70).
 CC -|- SUBUNIT: HOMODIMER, LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -|- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. STRONGEST
 CC SIMILARITY TO CTLA-4.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

EMBL; X93304; CAA63707.1; -;
 Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane.
 KW SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 219 T-CELL-SPECIFIC SURFACE GLYCOPROTEIN
 FT CD28.
 FT DOMAIN 19 151 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 152 178 POTENTIAL.
 FT DOMAIN 179 219 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 28 136 IG-LIKE V-TYPE DOMAIN.
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 219 AA; 25143 MW; 85B5C650E9634AA CRC64;

Query Match 13.0%; Score 141; DB 1; Length 219;
 Best Local Similarity 23.2%; Pred. No. 1e-06;
 Matches 36; Conservative 32; Mismatches 71; Indels 16; Gaps 6;

QY 30 MFIHNGGVOILCKYPD--IVQOFKMQLLKGGQILCDLTKTGSGNTVSISLKF-CHS 85
 Db 28 MLVYNNNEVNLCKTYTNLFSEKFRASLYKGADSAVEVCVNGNFSHPHFHSTTGFNCTV 87
 QY 86 QLSNNSVSFFLYNLDHSHANYFCNLISFDPPPF--KVTLTGGYLHIYESQLC----- 136

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:39:45 ; Search time 23.48 Seconds
(without alignments)
1121.325 Million cell updates/sec

Title: US-09-383-551B-2
Perfect score: 1082
Sequence: 1 MMSGWLWYFFLCRLIKVLGT.....YMFMRVNTAKSKSLTDTVTL 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

al number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1082	100.0	199	4 Q9Y6W8	Q9Y6W8 homo sapien
2	737.5	68.2	200	11 Q9WVS0	Q9WVS0 mus musculu
3	734	67.8	200	11 Q9JLV3	Q9JLV3 mus musculu
4	701	64.8	200	11 Q9RLT7	Q9RLT7 rattus norv
5	696	64.3	216	11 Q9WVR9	Q9WVR9 rattus norv
6	146.5	13.5	221	11 Q9JLV4	Q9JLV4 marmota mon
7	145.5	13.4	221	6 Q02757	Q02757 felis silve
8	145.5	13.4	221	6 Q9N214	Q9N214 felis silve
9	145	13.4	219	6 Q97630	Q97630 ovis aries
10	140.5	13.0	221	6 Q9N0N8	Q9N0N8 canis fami
11	140.5	13.0	221	6 Q9GKP3	Q9GKP3 canis fami
12	139.5	12.9	173	6 Q28289	Q28289 canis fami
13	103.5	9.6	988	5 Q17710	Q17710 caenorhabdi
14	91.5	8.5	485	5 Q20139	Q20139 caenorhabdi
15	89	8.2	321	4 Q9UGF6	Q9UGF6 homo sapien
16	87	8.0	563	5 Q17395	Q17395 caenorhabdi
17	84.5	7.8	209	4 Q9NYK4	Q9NYK4 homo sapien
18	84.5	7.8	223	11 Q9QZ27	Q9QZ27 mus musculu
19	84	7.8	285	1 Q28747	Q28747 archaeglob

20	84	7.8	419	13 Q9IA91	Q9IA91 morone saxa
21	83.5	7.7	223	11 Q9JLV3	Q9JLV3 marmota mon
22	83	7.7	1239	10 Q9FHM1	Q9FHM1 arabidopsis
23	82.5	7.6	223	11 Q62859	Q62859 rattus norv
24	81	7.5	276	14 Q98822	Q98822 human adeno
25	81	7.5	276	14 Q64861	Q64861 human adeno
26	81	7.5	680	11 Q55001	Q55001 mus musculu
27	81	7.5	1382	11 P97523	P97523 rattus norv
28	80.5	7.4	635	11 Q55002	Q55002 mus musculu
29	80.5	7.4	663	11 Q70434	Q70434 mus musculu
30	79.5	7.3	269	4 Q95297	Q95297 homo sapien
31	78.5	7.3	310	11 Q9EQ87	Q9EQ87 mus musculu
32	78	7.2	231	5 Q45688	Q45688 caenorhabdi
33	77.5	7.2	311	11 Q9EQ86	Q9EQ86 mus musculu
34	77	7.1	338	5 Q22576	Q22576 caenorhabdi
35	77	7.1	539	14 P88842	P88842 avian infec
36	77	7.1	619	10 Q9XG11	Q9XG11 sorghum bic
37	76.5	7.1	296	14 Q73315	Q73315 human immun
38	76.5	7.1	448	10 Q04353	Q04353 borago offi
39	76.5	7.1	1079	5 Q9VNP2	Q9VNP2 drosophila
40	76.5	7.1	1237	13 Q91976	Q91976 gallus gall
41	76.5	7.1	1475	5 Q25842	Q25842 plasmodium
42	76	7.0	174	11 Q921A7	Q921A7 rattus norv
43	76	7.0	394	11 Q9EQK9	Q9EQK9 rattus norv
44	76	7.0	425	6 Q9MZU2	Q9MZU2 oryctolagus
45	76	7.0	429	6 Q9MZU3	Q9MZU3 oryctolagus

ALIGNMENTS

RESULT 1

ID	Q9Y6W8	PRELIMINARY;	PRT;	199 AA.
AC	Q9Y6W8;			
DT	01-NOV-1999	(TREMBLrel. 12, Created)		
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)		
DE	ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM			
DE	PRECURSOR (INDUCIBLE T-CELL CO-STIMULATOR PRECURSOR).			
GN	ICOS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI-TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BLOOD;			
RA	Tezuka K., Tamatani T.;			
RT	"Cell surface molecule mediating cell adhesion and signal			
RT	transmission.";			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99127892; PubMed=9930702;			
RA	Hutloff A., Dittlich A.M., Beier K.C., Eljaschewitsch B., Kraft R.,			
RA	Anagnostopoulos I., Krocsek R.A.;			
RT	"ICOS is an inducible T-cell co-stimulator structurally and			
RT	functionally related to CD28.";			
RL	Nature 397:263-266(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Beier K.C., Hutloff A., Dittlich A.M., Heuck C., Mages H.W.,			
RA	Buecher K., Henn V., Rauch A., Krocsek R.A.;			
RT	"Detailed analysis of human ICOS and its ligand.";			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB023135; BAA82129.1; -.			
DR	EMBL; AJ277832; CAC06612.1; -.			
KW	Signal.			
FT	SIGNAL			
FT	1 20 POTENTIAL.			
SQ	SEQUENCE 199 AA; 22624 MW; 214EC741C9BDC9FC CRC64;			

```
Query Match 100.0%; Score 1082; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.8e-107;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKSGLWFFFLCRLIKVLTGEINGSANYEMFIHNGGVQLICKYPTDVQFKMQLKGGQ 60
Db 1 MKSGLWFFFLCRLIKVLTGEINGSANYEMFIHNGGVQLICKYPTDVQFKMQLKGGQ 60
QY 61 ILCDLTKTGSGNTVSTKSLKFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
Db 61 ILCDLTKTGSGNTVSTKSLKFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
QY 121 VLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCILICWLTKKKYSSSVHDPNGEY 180
Db 121 VLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCILICWLTKKKYSSSVHDPNGEY 180
QY 181 MFRAVNTAKKSLRLTDVTL 199
Db 181 MFRAVNTAKKSLRLTDVTL 199
RESULT 2
Q9WVS0 PRELIMINARY; PRT; 200 AA.
ID Q9WVS0
AC Q9WVS0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
DE PRECURSOR (CCLP PRECURSOR) (SURFACE PROTEIN).
GN CCCLP OR ICOS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINAL INTRA-EPITHELIUM;
RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Sarmiento U., Guo J.,
RA Horan T., Shih G., Zhang M., Coccia M.A., Kohno T., Tafuri-Bladt A.,
RA Campbell P., Chang D., Chiu L., Dai T., Duncan G., Elliott G.S.,
RA Hui A., McCabe S.M., Scully S., Shaklee C.L., Van G., Mak T.W.,
RA Senaldi G.;
RT "T-cell co-stimulation through B7RP-1 and ICOS.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RA McAdam A.J., Greenwald R.J., Levin M.A., Chernova T., Malenkovich N.,
RA Ling V., Freeman G.J., Sharpe A.H.;
RT "The inducible costimulatory (ICOS) molecule is critical for CD40-
RT mediated antibody class switching";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216748; AAF45150.1; -.
DR EMBL; AF327185; AAG48732.1; -.
DR EMBL; AF327184; AAG48732.1; JOINED.
FT NON_TER 200
SQ SEQUENCE 200 AA; 22709 MW; 87D97F0DC44ADCA7 CRC64;

Query Match 67.8%; Score 734; DB 11; Length 200;
Best Local Similarity 70.4%; Pred. No. 2.8e-70;
Matches 138; Conservative 19; Mismatches 35; Indels 4; Gaps 2;

QY 7 YF---FLFCLRIKVLTEINGSANYEMFIHNGGVQLICKYPTDVQFKMQLKGGQILC 63
Db 4 YFCRVFVFCFLIRLLTGEINGSADHRMFSFHNGGVQISCKYPETVQQLKMRLEFRERE 63
QY 64 DLTKTGSGNTVSTKSLKFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK-V 122
Db 64 ELTKTGSGNAVSIKPNMCLCYHLNNSVSFFLNNPDSSOGSYFFCSLSIFDPPPFERN 123
QY 123 LTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCILICWLTKKKYSSSVHDPNGEYMF 182
Db 124 LSGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCILICWLTKKKYSSSVHDPNGEYMF 183
QY 183 MRAVNTAKKSLRLTDVTL 198
Db 184 MAAVNTNKKSLAGVT 199
```

```
Query Match 100.0%; Score 1082; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.8e-107;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKSGLWFFFLCRLIKVLTGEINGSANYEMFIHNGGVQLICKYPTDVQFKMQLKGGQ 60
Db 1 MKSGLWFFFLCRLIKVLTGEINGSANYEMFIHNGGVQLICKYPTDVQFKMQLKGGQ 60
QY 61 ILCDLTKTGSGNTVSTKSLKFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
Db 61 ILCDLTKTGSGNTVSTKSLKFCHSOLSNNSVSFFLYNLDHSHANYFCNLSIFDPPPFK 120
QY 121 VLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCILICWLTKKKYSSSVHDPNGEY 180
Db 121 VLTGGYLHIYESQLCCQLKFWLPIGCAAFVVCILGCILICWLTKKKYSSSVHDPNGEY 180
QY 181 MFRAVNTAKKSLRLTDVTL 199
Db 181 MFRAVNTAKKSLRLTDVTL 199
RESULT 2
Q9WVS0 PRELIMINARY; PRT; 200 AA.
ID Q9WVS0
AC Q9WVS0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
DE PRECURSOR (CCLP PRECURSOR) (SURFACE PROTEIN).
GN CCCLP OR ICOS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Tezuka K., Tamatani T.;
RT "Cell surface molecule mediating cell adhesion and signal
RT transmission.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=SPLEEN;
RA Wu D., Glannoni M.A., Kisecker C.L., Faas S.J., Mickle A.P.,
RA Mattis L.A., Rother R.P.;
RT "cCLP, A novel molecule that regulates T cell activation.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=SPLEEN;
RX MEDLINE=20225659; PubMed=10760791;
RA Mages H.W., Hutloff A., Heuck C., Buchner K., Himmelbauer H.,
RA Oliveri F., Kroccek R.A.;
RT "Molecular cloning and characterization of murine ICOS and
RT identification of B7h as ICOS ligand.";
RL Eur. J. Immunol. 30:1040-1047(2000).
DR EMBL; AB023132; BAA82126.1; -.
DR EMBL; AF257230; AAF70099.1; -.
DR EMBL; AJ250559; CAB71153.1; -.
KW Signal.
FT SIGNAL 1 20 POTENTIAL.
SQ SEQUENCE 200 AA; 22690 MW; 9B2278E4CAB1DB47 CRC64;

Query Match 68.2%; Score 737.5; DB 11; Length 200;
Best Local Similarity 69.3%; Pred. No. 1.2e-70;
Matches 138; Conservative 20; Mismatches 40; Indels 1; Gaps 1;

QY 1 MKSGLWFFFLCRLIKVLTGEINGSANYEMFIHNGGVQLICKYPTDVQFKMQLKGGQ 60
Db 1 MKPYFCHVFCFLIRLLTGEINGSADHRMFSFHNGGVQISCKYPETVQQLKMRLEFRERE 60
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:46:03 ; Search time 32.53 Seconds
(without alignments)
14.050 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPDPF 6

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : PIR 68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6	100.0	199	2	S78540
2	6	100.0	200	2	JC7397
3	6	100.0	216	2	JC7396
4	5	83.3	41	2	PL0028
5	5	83.3	122	2	D70633
6	5	83.3	208	2	T02447
7	5	83.3	259	2	T29459
8	5	83.3	306	2	T49541
9	5	83.3	316	2	T25179
10	5	83.3	332	2	E82678
11	5	83.3	335	1	T02809
12	5	83.3	387	2	T22841
13	5	83.3	399	2	I49754
14	5	83.3	407	2	JC5851
15	5	83.3	408	2	S58174
16	5	83.3	409	2	S69229
17	5	83.3	425	2	T41683
18	5	83.3	438	2	T45823
19	5	83.3	441	2	H70632
20	5	83.3	449	2	T48199
21	5	83.3	468	1	F0WVMU
22	5	83.3	471	2	G72518
23	5	83.3	479	2	G86207
24	5	83.3	503	2	H70509
25	5	83.3	518	1	S27381
26	5	83.3	538	1	F0WVIM
27	5	83.3	538	1	F0WVM
28	5	83.3	538	1	Q0ECD7
29	5	83.3	585	2	S74673

30	5	83.3	695	2	E82783	hypothetical prote
31	5	83.3	720	2	E72074	1,4-alpha-glucan b
32	5	83.3	720	2	G86549	glucan branching e
33	5	83.3	770	2	B56695	transducin-like en
34	5	83.3	793	1	JC6161	kinesin-associated
35	5	83.3	852	2	T33824	hypothetical prote
36	5	83.3	877	1	A25962	mitosis inhibitor
37	5	83.3	922	2	D75615	exonuclease ABC c
38	5	83.3	981	1	F0WVGM	gag-abl polyprotei
39	5	83.3	1116	2	T16112	hypothetical prote
40	5	83.3	1183	2	S65236	probable membrane
41	5	83.3	1190	2	T38636	tat binding homolo
42	5	83.3	1255	1	A37967	neural cell adhesi
43	5	83.3	1575	2	S68448	synaptotagmin, 170K
44	5	83.3	3027	2	JQ1917	polyprotein - pars
45	4	66.7	11	2	D45900	complement C3b rec
46	4	66.7	27	2	PL0029	plasma protein Po2
47	4	66.7	28	2	T14905	hypothetical prote
48	4	66.7	35	2	T09834	hypothetical prote
49	4	66.7	40	2	PL0030	plasma protein Po2
50	4	66.7	46	2	T07459	hypothetical prote

ALIGNMENTS

RESULT 1

S78540

inducible T-cell co-stimulator ICOS precursor - human

C:Species: Homo sapiens (man)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 07-May-1999

C:Accession: S78540; S78748; S78749

R:Krocze, R.

submitted to the Protein Sequence Database, June 1998

A:Reference number: S78540

A:Accession: S78540

A:Molecule type: mRNA

A:Residues: 1-199 <KRO>

A:Experimental source: cell line MOLT-4V

R:Hutloff, A.; Dittich, A.M.; Beier, K.C.; Eljaschewitsch, B.; Kraft, R.; Anagnostop

Nature 397, 263-266, 1999

A:Title: ICOS is an inducible T-cell co-stimulator structurally and functionally rela

A:Reference number: S78748; MUID:99127892

A:Accession: S78748

A:Molecule type: mRNA

A:Residues: 1-199 <HUT1>

A:Experimental source: cell line MOLT-4V

A:Accession: S78749

A:Molecule type: protein

A:Residues: 'X',193-198 <HUT2>

A:Experimental source: cell line MOLT-4V

C:Complex: homodimer

C:Superfamily: immunoglobulin homology

C:Keywords: dimer; glycoprotein; T-cell; transmembrane protein

F:1-19/Domain: (or 1-20) signal sequence #status predicted <SIG>

F:20-199/Product: (or 21-199) inducible costimulator ICOS #status predicted <MAT>

F:21-138/Domain: extracellular #status predicted <EXT>

F:26-132/Domain: immunoglobulin homology <IMM>

F:139-164/Domain: transmembrane #status predicted <TM>

F:165-199/Domain: intracellular #status predicted <INT>

F:23,89,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 6; DB 2; Length 199;

Best Local Similarity 100.0%; Pred. No. 4.5;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6

|||||

Db 114 FDPDPF 119

RESULT 2

JC7397
activation-inducible lymphocyte immuno-mediatory molecule-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7397; PC7100
R:Tezuka, K.; Tsujil, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kamada, Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat ALLIM/ICOS, a novel T-cell costimulatory molecule
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7397
A:Molecule type: mRNA
A:Residues: 1-200 <TE2>
A:Cross-references: DDBJ:AB023134
A:Accession: PC7100
A:Molecule type: protein
A:Residues: 21-40 <TE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
C:Gene: allIm-2
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 6; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
|||||
Db 114 FDPDPF 119

RESULT 3
JC7396
activation-inducible lymphocyte immuno-mediatory molecule-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7396; PC7099
R:Tezuka, K.; Tsujil, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kamada, Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat ALLIM/ICOS, a novel T-cell costimulatory molecule
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7396
A:Molecule type: mRNA
A:Residues: 1-216 <TE2>
A:Cross-references: DDBJ:AB023133
A:Accession: PC7099
A:Molecule type: protein
A:Residues: 21-40 <TE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
C:Gene: allIm-1
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 6; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
|||||
Db 114 FDPDPF 119

RESULT 4
PL0028
plasma protein Xk - horse (fragment)
N:Alternate names: alpha-1-glycoprotein homolog
C:Species: Equus caballus (domestic horse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 11-May-2000

C:Accession: PL0028
R:Van de Weghe, A.; Coppieters, W.; Bauw, G.; Vandekerckhove, J.; Bouquet, Y. Comp. Biochem. Physiol. B 90, 751-756, 1988
A:Title: The homology between the serum proteins Po2 in pig, Xk in horse and alpha 1B
A:Reference number: PL0028; MUID:89250430
A:Accession: PL0028
A:Molecule type: protein
A:Residues: 1-41 <VAN>
A:Experimental source: serum
A:Note: 11-Leu, 26-Val, and 37-Glu were also found
C:Keywords: glycoprotein; plasma

Query Match 83.3%; Score 5; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 5
|||||
Db 4 FDPDPF 8

RESULT 5
D70633
hypothetical protein RV0397 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70633
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Spates, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: D70633
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-122 <COL>
A:Cross-references: GB:Z84725; GB:AL123456; NID:g3261703; PIDN:CA806603.1; PID:e30041
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0397

Query Match 83.3%; Score 5; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDPF 6
|||||
Db 118 DPDPF 122

RESULT 6
T02447
hypothetical protein At2g46000 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F41l8.2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02447; B84897
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome II BAC F41l8 genomic sequence.
A:Reference number: Z14674
A:Accession: T02447
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-208 <ROU>
A:Cross-references: EMBL:AC004665; NID:g3386593; PID:g3386595
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:48 ; Search time 17.42 Seconds
(without alignments)
11.799 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPPPP 6

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5	83.3	41	1	ALBG_HORSE
2	5	83.3	235	1	GAG_MLVAB
3	5	83.3	343	1	BBUR_BORBR
4	5	83.3	399	1	HXAA_MOUSE
5	5	83.3	468	1	GAG_MSVMT
6	5	83.3	518	1	KKL6_YEAST
7	5	83.3	536	1	GAG_MLVCB
8	5	83.3	538	1	GAG_MLVMO
9	5	83.3	538	1	GAG_MSVMO
10	5	83.3	538	1	TNSE_ECOLI
11	5	83.3	770	1	TLEI_HUMAN
12	5	83.3	770	1	TLEI_MOUSE
13	5	83.3	877	1	WEEL_SCHPO
14	5	83.3	1121	1	DDX8_ARATH
15	5	83.3	1266	1	NGCA_CHICK
16	5	83.3	1324	1	SYU1_BOVIN
17	5	83.3	1574	1	SYJ1_RAT
18	5	83.3	1575	1	SYJ1_HUMAN
19	5	83.3	3027	1	POLG_PYFV1
20	4	66.7	40	1	ALBG_PIG
21	4	66.7	60	1	PHX1_MOUSE
22	4	66.7	60	1	Y01E_BPT4
23	4	66.7	71	1	FBS_MOUSE
24	4	66.7	82	1	PN3A_PENVA
25	4	66.7	82	1	PN3B_PENVA
26	4	66.7	91	1	YPD1_ECOLI
27	4	66.7	92	1	Y676_TREPA
28	4	66.7	94	1	VAI5_VACCC
29	4	66.7	94	1	VAI5_VARY
30	4	66.7	106	1	B2_BBV
31	4	66.7	115	1	IPSG_VULVU
32	4	66.7	118	1	REV_HVILW
33	4	66.7	119	1	HYP3_AGABI

34	4	66.7	129	1	VG58_BPMD2
35	4	66.7	137	1	V137_ASFL5
36	4	66.7	139	1	VG26_HSVII
37	4	66.7	143	1	NIFX_FRAAL
38	4	66.7	144	1	INDC_BOVIN
39	4	66.7	146	1	RPOB_LIBAF
40	4	66.7	146	1	YN81_YEAST
41	4	66.7	151	1	YWIS_WHEAT
42	4	66.7	157	1	VES_RHPV1
43	4	66.7	158	1	MB27_BOVIN
44	4	66.7	159	1	COAD_TREPA
45	4	66.7	159	1	MB28_BOVIN
46	4	66.7	159	1	NIFX_RHOCA
47	4	66.7	160	1	SC51_SHEEP
48	4	66.7	160	1	SC52_SHEEP
49	4	66.7	161	1	COAD_AQUAE
50	4	66.7	162	1	SKPL_HUMAN

ALIGNMENTS

RESULT	1
ALBG_HORSE	
ID	ALBG_HORSE
AC	P3091;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-FEB-1995 (Rel. 31, Last sequence update)
DT	01-FEB-1995 (Rel. 31, Last annotation update)
DE	ALPHA-1B-GLYCOPROTEIN (PLASMA PROTEIN XK) (FRAGMENT).
OS	Equus caballus (Horse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX	NCBI_TaxID=9796;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Plasma;
RX	MEDLINE=89250430; PubMed=3248368;
RA	van de Weghe A., Coppleters W., Bauw G., Vandekerckhove J.,
RA	Bouquet Y.;
RT	"The homology between the serum proteins PO2 in pig, Xk in horse and
RT	alpha 1B-glycoprotein in human."
RL	Comp. Biochem. Physiol. 90B:751-756(1988).
CC	-!- FUNCTION: PLASMA GLYCOPROTEIN OF UNKNOWN FUNCTION.
CC	-!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
DR	PIR; P10028; P10028.
KW	Immunoglobulin domain; Glycoprotein; Plasma.
FT	VARIANT 11 11 W -> L.
FT	VARIANT 26 26 Q -> V.
FT	VARIANT 37 37 G -> E.
FT	NON_TER 41 41
SQ	SEQUENCE 41 AA; 4457 MW; 2AA928CBA51FFED6 CRC64;

Query Match 83.3%; Score 5; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FDPPPP 5
Db	4	FDPPPP 8

RESULT	2
GAG_MLVAB	
ID	GAG_MLVAB
AC	P03333;
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	01-FEB-1991 (Rel. 17, Last annotation update)
DE	GAG POLYPROTEIN [CONTAINS: CORE PROTEIN P15; INNER COAT PROTEIN P12; CORE SHELL PROTEIN P30].

GN . GAG.
 OS Abelson murine leukemia virus.
 OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_TaxID=11788;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=83221648; PubMed=6304726;
 RX Reddy E.P., Smith M.J., Srinivasan A.;
 RT "Nucleotide sequence of Abelson murine leukemia virus genome:
 RT structural similarity of its transforming gene product to other onc
 RT gene products with tyrosine-specific kinase activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).
 CC -1- PTM: SPECIFIC ENZYMIC CLAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL
 CC POLYPEPTIDE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; V01541; CAA24781.1; -;
 DR PIR; A03931; FOMVGM.
 DR InterPro; IPR000840; -;
 DR InterPro; IPR002079; -;
 DR InterPro; IPR003036; -;
 DR Pfam; PF02093; Gag_p30; 1.
 DR Pfam; PF01140; gag_MA; 1.
 DR Pfam; PF01141; gag_p12; 1.
 KW Coat protein; Core protein; Polyprotein; Myristate.
 FT CHAIN 2 131 CORE PROTEIN P15.
 FT CHAIN 132 215 INNER COAT PROTEIN P12.
 FT CHAIN 216 235 CORE SHELL PROTEIN P30.
 FT LIPID 2 2 MYRISTATE.
 SQ SEQUENCE 235 AA; 25641 MW; 4D83FF1D7E056C7D CRC64;
 Query Match 83.3%; Score 5; DB 1; Length 235;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FDPDP 5
 Db 91 FDPDP 95
 RESULT 3
 BBUR_BBUR
 ID BBUR_BBUR STANDARD; PRT; 343 AA.
 AC 006703;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROBABLE TRANSCRIPTIONAL REGULATOR BBUR.
 GN BBUR.
 OS Bordetella bronchiseptica.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BB7866;
 RX McMillan D.J., Mau M., Walker M.J.;
 RT "Characterisation of the urease gene cluster in Bordetella
 RT bronchiseptica";
 RL Gene 208:243-251(1998).
 CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF000579; AAC46123.1; -;
 DR InterPro; IPR000847; -;
 DR Pfam; PF00126; HTH_1; 1.
 DR PRINTS; PRO0039; HTHLYSR.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 KW Transcription regulation; DNA-binding.
 FT DNA_BIND 32 51 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 343 AA; 36629 MW; 93437DFC522CDF6F CRC64;
 Query Match 83.3%; Score 5; DB 1; Length 343;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FDPDP 5
 Db 161 FDPDP 165
 RESULT 4
 HXAA_MOUSE
 ID HXAA_MOUSE STANDARD; PRT; 399 AA.
 AC P13110;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HOMEOBOX PROTEIN HOX-A10 (HOX-1.8).
 GN HOXA10 OR HOXA-10 OR HOX-1.8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Kidney;
 RX MEDLINE=95166244; PubMed=7862151;
 RA Benson G.V., Nguyen T.-H.E., Maas R.L.;
 RT "The expression pattern of the murine Hoxa-10 gene and the sequence
 RT recognition of its homeodomain reveal specific properties of
 RT Abdominal B-like genes";
 RL Mol. Cell. Biol. 15:1591-1601(1995).
 RN [2]
 RP SEQUENCE OF 325-384 FROM N.A.
 RX MEDLINE=92073356; PubMed=1683707;
 RA Singh G., Kaur S., Stock J.L., Jenkins N.A., Gilbert D.J.,
 RA Copeland N.G., Potter S.S.;
 RT "Identification of 10 murine homeobox genes";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10706-10710(1991).
 RN [3]
 RP SEQUENCE OF 346-370 FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Spleen;
 RX MEDLINE=92073357; PubMed=1720547;
 RA Murtha M.T., Leckman J.F., Ruddle F.H.;
 RT "Detection of homeobox genes in development and evolution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715(1991).
 CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
 CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
 CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
 CC BINDS TO THE DNA SEQUENCE 5'-AA[AT]TTTATTAC-3'.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING LIMB BUD WHERE IT
 CC IS RESTRICTED TO THE MESENCHYME ALONG THE PROXIMAL-DISTAL AXIS.
 CC ALSO FOUND IN DEVELOPING GUT AND UROGENITAL TRACT. IN ADULT

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:19 ; Search time 47.57 Seconds
(without alignments)
16.688 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPFPF 6

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6	100.0	199	4 Q9Y6W8	Q9Y6W8 homo sapien
2	6	100.0	200	11 Q9R1T7	Q9R1T7 ratus norv
3	6	100.0	200	11 Q9WVS0	Q9WVS0 mus musculu
4	6	100.0	200	11 Q9JL17	Q9JL17 mus musculu
5	6	100.0	216	11 Q9WVR9	Q9WVR9 ratus norv
6	5	83.3	121	11 P70634	P70634 ratus norv
7	5	83.3	122	2 P95205	P95205 mycobacteri
8	5	83.3	143	14 O71110	O71110 bovine aden
9	5	83.3	208	10 O82812	O82812 arabidopsis
10	5	83.3	220	3 Q9U0U0	Q9U0U0 schizosacch
11	5	83.3	259	2 O86634	O86634 streptomyce
12	5	83.3	306	3 Q9P628	Q9P628 neurospora
13	5	83.3	316	5 Q22696	Q22696 caenorhabdi
14	5	83.3	332	2 Q9PDB5	Q9PDB5 xylella fas
15	5	83.3	335	5 Q25340	Q25340 leishmania
16	5	83.3	371	5 Q9GY95	Q9GY95 leishmania
17	5	83.3	382	5 Q9NGK0	Q9NGK0 leishmania
18	5	83.3	387	5 O62265	O62265 caenorhabdi
19	5	83.3	398	2 Q55224	Q55224 streptomyce

20	5	83.3	407	2	O32452	O32452 actinomadr
21	5	83.3	408	2	O53672	O53672 streptomyce
22	5	83.3	409	2	O54496	O54496 streptomyce
23	5	83.3	415	2	O9F6E1	O9F6E1 streptomyce
24	5	83.3	425	2	O54815	O54815 streptomyce
25	5	83.3	425	3	Q9U0U3	Q9U0U3 schizosacch
26	5	83.3	438	10	Q9M3B4	Q9M3B4 arabidopsis
27	5	83.3	441	2	P95201	P95201 mycobacteri
28	5	83.3	449	10	Q9LZW8	Q9LZW8 arabidopsis
29	5	83.3	460	2	Q9XZD9	Q9XZD9 actinomadr
30	5	83.3	471	1	O9YA15	O9YA15 aeropyrum p
31	5	83.3	479	10	Q9LML6	Q9LML6 arabidopsis
32	5	83.3	498	14	O9WNA9	O9WNA9 kaposi's sa
33	5	83.3	503	2	O33266	O33266 mycobacteri
34	5	83.3	503	2	O33360	O33360 mycobacteri
35	5	83.3	534	14	Q9YKA0	Q9YKA0 murine leuk
36	5	83.3	537	14	O56222	O56222 moloney mur
37	5	83.3	537	14	Q9WJP4	Q9WJP4 moloney mur
38	5	83.3	552	14	O69102	O69102 herpes simp
39	5	83.3	585	2	P72809	P72809 synecocyst
40	5	83.3	662	5	Q9W4S0	Q9W4S0 drosophila
41	5	83.3	695	2	O9PFP3	O9PFP3 xylella fas
42	5	83.3	708	5	O9GVA7	O9GVA7 pimla hypo
43	5	83.3	720	2	O9Z876	O9Z876 chlamydia p
44	5	83.3	722	5	Q9U0Z4	Q9U0Z4 leishmania
45	5	83.3	792	4	Q92845	Q92845 homo sapien
46	5	83.3	793	11	P70188	P70188 mus musculu
47	5	83.3	852	5	O9TYT5	O9TYT5 caenorhabdi
48	5	83.3	922	2	Q9RYW8	Q9RYW8 deinococcus
49	5	83.3	980	14	O92809	O92809 abelson mur
50	5	83.3	981	14	O39477	O39477 abelson mur

ALIGNMENTS

RESULT 1
Q9Y6W8 PRELIMINARY; PRT; 199 AA.
ID Q9Y6W8;
AC Q9Y6W8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE ALLIM
DE PRECURSOR (INDUCIBLE T-CELL CO-STIMULATOR PRECURSOR).
GN ICOS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Tezuka K., Tamatani T.;
RT "Cell surface molecule mediating cell adhesion and signal
transmission.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99127892; PubMed=9930702;
RA Hutloff A., Dittrich A.M., Beier K.C., Eljaschewitsch B., Kraft R.,
Anagnostopoulos I., Kroccek R.A.;
RT "ICOS is an inducible T-cell co-stimulator structurally and
functionally related to CD28.";
RL Nature 397:263-266(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Beier K.C., Hutloff A., Dittrich A.M., Heuck C., Mages H.W.,
Buechner K., Henn V., Rauch A., Kroccek R.A.;
RT "Detailed analysis of human ICOS and its ligand.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023135; BAA82129.1; -
DR EMBL; AJ277832; CAC06612.1; -

```
KW . Signal.
FT SIGNAL 1 20 POTENTIAL.
SQ SEQUENCE 199 AA; 22624 MW; 214EC741C9BDC9FC CRC64;

. Query Match 100.0%; Score 6; DB 4; Length 199;
  Best Local Similarity 100.0%; Pred. No. 3.5;
  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 FDPDPF 119

RESULT 2
Q9RLT7 PRELIMINARY; PRT; 200 AA.
AC Q9RLT7
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DR 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Tezuka K., Tamatani T.;
RT "Cell surface molecule mediating cell adhesion and signal
  transmission.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023134; BAA82128.1; -.
KW Signal.
FT SIGNAL 1 20 POTENTIAL.
SQ SEQUENCE 200 AA; 22529 MW; 0A74C35581F129D4 CRC64;

Query Match 100.0%; Score 6; DB 11; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 FDPDPF 119

LT 3
SQ SEQUENCE FROM N.A.
ID Q9WVS0 PRELIMINARY; PRT; 200 AA.
AC Q9WVS0
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DR 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE ACTIVATION-INDUCIBLE LYMPHOCYTE IMMUNOMEDIATORY MOLECULE AILIM
DE PRECURSOR (CCLP PRECURSOR) (SURFACE PROTEIN).
GN CCPLP OR ICOS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RA Tezuka K., Tamatani T.;
RT "Cell surface molecule mediating cell adhesion and signal
  transmission.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-SPLEEN;
RX MEDLINE=20225659; PubMed=10760791;
RA Mages H.W., Hutloff A., Heuck C., Buchner K., Himmelbauer H.,
  Oliveri F., Krocsek R.A.;
RT "Molecular cloning and characterization of murine ICOS and
  identification of B7h as ICOS ligand.";
RL Eur. J. Immunol. 30:1040-1047(2000).
DR EMBL; AB023132; BAA82126.1; -.
DR EMBL; AF257230; AAF70099.1; -.
DR EMBL; AJ250559; CAB71153.1; -.
KW Signal.
FT SIGNAL 1 20 POTENTIAL.
SQ SEQUENCE 200 AA; 22690 MW; 9B2278E4CAB1DB47 CRC64;

Query Match 100.0%; Score 6; DB 11; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 FDPDPF 119

RESULT 4
Q9JL17 PRELIMINARY; PRT; 200 AA.
ID Q9JL17
AC Q9JL17
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DR 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE CD28-RELATED PROTEIN 1 (INDUCIBLE COSTIMULATORY PROTEIN) (FRAGMENT).
OS ICOS.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINAL INTRA-EPITHELIUM;
RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Sarmiento U., Guo J.,
  Horan T., Shih G., Zhang M., Coccia M.A., Kohno T., Tafuri-Bladt A.,
  Campbell P., Chang D., Chiu L., Dai T., Duncan G., Elliott G.S.,
  Hui A., McCabe S.M., Scully S., Shaklee C.L., Van G., Mak T.W.,
  Senaldi G.;
RT "T-cell co-stimulation through B7RP-1 and ICOS.";
RL Nature 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129/OLA;
RA McAdam A.J., Greenwald R.J., Levin M.A., Chernova T., Malenkovich N.,
  Ling V., Freeman G.J., Sharpe A.H.;
RT "The inducible costimulatory (ICOS) molecule is critical for CD40-
  mediated antibody class switching.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216748; AAF45150.1; -.
DR EMBL; AF327185; AAG48732.1; -.
DR EMBL; AF327184; AAG48732.1; JOINED.
FT NON_TER 200 200
SQ SEQUENCE 200 AA; 22709 MW; 87D97F0DC44ADCA7 CRC64;

Query Match 100.0%; Score 6; DB 11; Length 200;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 FDPDPF 119
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:44:39 ; Search time 45.63 Seconds
(without alignments)
7.972 Million cell updates/sec

Title: US-09-383-551b-21

Perfect score: 6

Sequence: 1 FDPFPF 6

Scoring table:

OLIGO Gapex 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	100.0	198	21	Human Th2-specific
2	6	100.0	199	19	Human cell surface
3	6	100.0	199	19	Human cell surface
4	6	100.0	199	20	Human activated T-
5	6	100.0	199	21	Amino acid sequenc
6	6	100.0	200	19	Mouse cell surface
7	6	100.0	200	19	Rat cell surface p
8	6	100.0	200	21	Amino acid sequenc
9	6	100.0	200	21	Murine Th2-specifi
10	6	100.0	216	19	Rat cell surface p
11	5	83.3	10	9	Hook region #5 con

12	5	83.3	13	9	AA80875	Hook region #6 con
13	5	83.3	20	9	AA80876	Hook region #7 con
14	5	83.3	87	21	AA840437	Human ORFX ORF201
15	5	83.3	89	22	AA863385	Human breast cance
16	5	83.3	131	16	AA876596	MOMLV mutated gag
17	5	83.3	131	16	AA876593	MOMLV gag matrix p
18	5	83.3	131	16	AA876595	MOMLV mutated gag
19	5	83.3	140	16	AA877751	MOMLV gag/pol wobb
20	5	83.3	140	17	AA877099	MOMLV gag gene prod
21	5	83.3	140	17	AA878101	Gag gene product.
22	5	83.3	140	21	AA878788	Moloney murine leu
23	5	83.3	141	18	AA878787	Representative "wo
24	5	83.3	141	21	AA878787	Moloney murine leu
25	5	83.3	189	21	AA815562	Arabidopsis thalia
26	5	83.3	189	21	AA859366	Arabidopsis thalia
27	5	83.3	208	21	AA815561	Arabidopsis thalia
28	5	83.3	208	21	AA859365	Arabidopsis thalia
29	5	83.3	234	21	AA826477	Arabidopsis thalia
30	5	83.3	265	21	AA807499	Arabidopsis thalia
31	5	83.3	265	21	AA853871	Arabidopsis thalia
32	5	83.3	345	21	AA878821	Amino acid sequenc
33	5	83.3	398	21	AA877269	Streptomyces sp. C
34	5	83.3	403	20	AA826882	Streptomyces sp. t
35	5	83.3	406	20	AA826880	Streptomyces sp. t
36	5	83.3	407	19	AA854381	Actinomadura hibis
37	5	83.3	409	17	AA805776	Streptomyces nogal
38	5	83.3	409	21	AA877274	Streptomyces negal
39	5	83.3	409	21	AA878826	Amino acid sequenc
40	5	83.3	425	21	AA877270	Streptomyces peuce
41	5	83.3	425	21	AA878822	Amino acid sequenc
42	5	83.3	438	21	AA807498	Arabidopsis thalia
43	5	83.3	438	21	AA853870	Arabidopsis thalia
44	5	83.3	538	21	AA810043	MOMLV gag protein.
45	5	83.3	648	20	AA817946	MOMLV gag gene prot
46	5	83.3	722	20	AA835095	C. pneumoniae prot
47	5	83.3	770	15	AA851109	Human TLE-1 protei
48	5	83.3	770	15	AA851476	Human TLE-1. Homo
49	5	83.3	1737	21	AA810044	MOMLV gag-pol prote
50	4	66.7	6	19	AA831462	Transcriptional ac

ALIGNMENTS

RESULT 1

AA892213

ID AA892213 standard; Protein; 198 AA.

XX AA892213;

AC AA892213;

XX 10-AUG-2000 (first entry)

XX Human Th2-specific polypeptide, h1228.

DE h1228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;

XX h1228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;

KW CD28; antidiabetic; immunosuppressive; neuroprotective; dermatological;

KW antipsoriatic; antiallergic; anti-viral; ophthalmological;

KW CUTA-4; nephrotropic; anti-HIV; antibacterial.

XX Homo sapiens.

OS Homo sapiens.

XX Key

FT Peptide

FT 113..118

FT /label= conserved_ppp_motif

FT /note= "common in CD28 and CTLA-4"

FT Peptide

FT 178..181

FT /label= YXXM_motif

FT /note= "common in CD28 and CTLA-4; necessary for

CD28-mediated phosphatidylinositol 3-kinase

XX activity"

XX WO200019988-A1.

XX

13-APR-2000.

06-OCT-1999; 99WO-US23156.

07-OCT-1998; 98US-0168229.

26-FEB-1999; 99US-0258670.

06-OCT-1999; 99US-04113136.

(MILL-) MILLENNIUM PHARM INC.

Lehar S, Manning S, Coyle AJ, Gutierrez-ramos J;

WPI; 2000-303619/26.

N-PSDB; AAA09056.

T helper (Th) 2 nucleic acids and encoded proteins, useful for the diagnosis and treatment of immune and respiratory disorders such as Crohn's diseases, arthritis, insulin dependent diabetes and autoimmunity

Claim 16; Page 138-139; 159pp; English.

This Th2-specific polypeptide, which has similarity to human CD28 and human CTLA-4 is encoded by human orthologue h128.

A novel method for modulating a Th2 response, an immune response, or suppressing airway inflammation or hyperresponsiveness in a mammal comprises administering a Th2-specific polypeptide of the invention, an antibody to such a polypeptide or allelic variants of the genes. The novel DNA and polypeptide sequences are useful for treatment and diagnosis of Th cell and Th cell-like disorders such as Crohn's diseases, arthritis, insulin dependent diabetes, autoimmunity, multiple sclerosis, dermatitis, psoriasis, graft rejection, graft versus host diseases, sarcoidosis, asthma, allergies, conjunctivitis, glomerular nephritis, viral infections (including human immunodeficiency virus (HIV)), bacterial infections, bronchitis, cystic fibrosis, diptheria, emphysema, pneumonia, and Legionnaires disease.

Sequence 198 AA;

Query Match 100.0%; Score 6; DB 21; Length 198;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
Db 114 fdpppf 119

AAW75956
AAW75956 standard; Protein; 199 AA.

AAW75956;

11-DEC-1998 (first entry)

Human cell surface protein #1.

Human; cell surface protein; thymocyte; lymphocyte; cell adhesion; signal transmission; autoimmune disorder; allergy; diagnosis; mitogen-stimulated.

Homo sapiens.

WO9838216-A1.

03-SEP-1998.

27-FEB-1998; 98WO-JF00837.

26-FEB-1998; 98JP-0062217.

27-SEP-1997; 97JP-0062290.

PS Claim 9; Page 101-105; 149pp; Japanese.

CC The present sequence represents a human cell surface protein which is
 CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
 CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
 CC antibodies recognising the cell surface protein. These antibodies also
 CC produce an increase in peripheral blood lymphocytes in the presence of
 CC an antibody recognising CD3 antigen. The cell surface protein contains
 CC the amino acid sequence FDPDPF in its extracellular region and the
 CC sequence YMF in its intracellular region. The cell surface protein can
 CC be used in the prevention and treatment of autoimmune and allergic
 CC diseases, and in the diagnosis and investigation of such disorders.

XX Sequence 199 AA;

Query Match 100.0%; Score 6; DB 19; Length 199;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FDPDPF 6

Db 114 fdpppf 119

RESULT 4

AA08026
 ID AAY08026 standard; Protein; 199 AA.

XX AC AAY08026;

XX DT 08-JUL-1999 (first entry)

XX DE Human activated T-lymphocyte protein 8F4.

XX KW T-lymphocyte; human; 8F4; T cell co-stimulation; activated; CD4+; CD8+;
 KW anticancer; antiviral; anti-asthma; immunomodulator; proliferation;
 KW T cell activation; cytokine synthesis; regulatory element; B cell;
 KW T cell-dependent antibody production; treatment; prevention; cancer;
 KW autoimmune disease; transplant rejection; immune system regulation;
 KW disorder; acquired immune deficiency syndrome; AIDS; asthma.

XX OS Homo sapiens.

XX PN WO915553-A2.

XX PD 01-APR-1999.

XX 23-SEP-1998; 98WO-DE02896.

PR 11-MAY-1998; 98DE-1021060.

PR 23-SEP-1997; 97DE-1041929.

(DEKO-) DEUT KOCH INST ROBERT.

(BUND) BUNDESREPUBLIK DEUT PAUL-EHRLICH-INST.

XX Krocsek R;

XX WPI; 1999-276975/23.

DR N-PSDB; AAX37661.

XX Polypeptide 8F4 co-stimulates T cells and is present only on
 PT activated cells

XX Claim 2; Page 24; 47pp; German.

XX This invention describes a novel human protein, 8F4, and its encoding
 CC nucleic acid which co-stimulates T cells and is present on activated CD4+
 CC and CD8+ T cells but not on resting or activated B cells, granulocytes,
 CC monocytes, natural killer or dendritic cells. 8F4 has anticancer,
 CC antiviral, anti-asthma and immunomodulatory activity. 8F4 provides a
 CC strong co-stimulatory signal for T cell activation, i.e. it amplifies
 CC proliferation of T cells, synthesis of certain cytokines and other

CC regulatory agents, and improves T cell-dependent antibody production
 CC by B cells. Agents that inhibit 8F4 are used to treat or prevent
 CC autoimmune diseases, to prevent transplant rejection and to treat
 CC disorders of immune system regulation. 8F4, or cells that express it,
 CC is/are used to treat or prevent cancers, acquired immune deficiency
 CC syndrome, asthma and chronic infectious diseases (e.g. hepatitis B or C).

SQ Sequence 199 AA;

Query Match 100.0%; Score 6; DB 20; Length 199;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6

Db 114 fdpppf 119

RESULT 5

AA080731
 ID AAB080731 standard; Protein; 199 AA.

XX AC AAB080731;

XX DT 02-JAN-2001 (first entry)

XX DE Amino acid sequence of a human CRP1 polypeptide.

XX KW CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
 KW T-lymphocyte activation; type I transmembrane protein; T cell activation;
 KW T cell proliferation; T-cell mediated disorder.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note= "signal peptide"

FT Domain /note= "mature protein"

FT Domain /note= "extracellular domain"

FT Domain /note= "predicted transmembrane domain"

FT Domain /note= "intracellular domain"

XX PN WO200046240-A2.

XX PD 10-AUG-2000.

XX 27-JAN-2000; 2000WO-US01871.

XX 03-FEB-1999; 99US-0244448.

XX 08-MAR-1999; 99US-0264527.

XX (AMGE-) AMGEN INC.

XX Yoshinaga SK;

XX WPI; 2000-543476/49.

XX N-PSDB; AAA64558.

XX Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
 PT in the treatment, prevention and diagnosis of T cell mediated disorders
 PT Disclosure; Fig 13A; 174pp; English.

XX The present sequence represents a CRP1 (CD28 related protein-1)
 CC polypeptide. The specification also describes a B7RP1 (B7 related
 CC protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
 CC activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are

CC . predicted to be a type I transmembrane protein. The nucleic acids are
CC useful for regulating T cell activation or proliferation in an animal.
CC The polypeptides are useful for treating, preventing ameliorating or
CC diagnosing a T-cell mediated disorder in an animal. They can also be
CC used to identify test molecules that bind to the polypeptides.
XX
SQ Sequence 199 AA;

Query Match 100.0%; Score 6; DB 21; Length 199;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
| | | | |
Db 114 fdpppf 119

RESULT 6
AAW75958
AAW75958 standard; Protein; 200 AA.

XX AC AAW75958;
XX DT 11-DEC-1998 (first entry)
XX DE Mouse cell surface protein.
XX KW Mouse; cell surface protein; thymocyte; lymphocyte; cell adhesion;
XX KW signal transmission; autoimmune disorder; allergy; diagnosis;
XX KW mitogen-stimulated.
XX OS Mus sp.
XX PN WO9838216-A1.
XX PD 03-SEP-1998.
XX PF 27-FEB-1998; 98WO-JP00837.
XX PR 26-FEB-1998; 98JP-0062217.
XX PR 27-FEB-1997; 97JP-0062290.
XX PA (NISR) JAPAN TOBACCO INC.
XX PI Tamatani T, Tezuka K;
XX WPI; 1998-481144/41.
XX N-PSDB; AAV53200.

PT Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to
PT it useful in treatment of autoimmune and allergic disorders.

PS Claim 9; Page 110-112; 149pp; Japanese.

XX The present sequence represents a mouse cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FDPDPF in its extracellular region and the
CC sequence YWFM in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.
XX
SQ Sequence 200 AA;

Query Match 100.0%; Score 6; DB 19; Length 200;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
| | | | |
Db 114 fdpppf 119

RESULT 7
AAW71874
AAW71874 standard; Protein; 200 AA.

XX AC AAW71874;
XX DT 11-DEC-1998 (first entry)
XX DE Rat cell surface protein #1.
XX KW Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
XX KW signal transmission; autoimmune disorder; allergy; diagnosis;
XX KW mitogen-stimulated.

XX OS Rattus sp.

XX PN WO9838216-A1.

XX PD 03-SEP-1998.

XX PF 27-FEB-1998; 98WO-JP00837.

XX PR 26-FEB-1998; 98JP-0062217.

XX PR 27-FEB-1997; 97JP-0062290.

XX PA (NISB) JAPAN TOBACCO INC.

XX PI Tamatani T, Tezuka K;

XX WPI; 1998-481144/41.

XX N-PSDB; AAV61357.

PT Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transmission and cell adhesion, and antibodies to
PT it useful in treatment of autoimmune and allergic disorders.

PS Claim 9; Page 106-109; 149pp; Japanese.

XX The present sequence represents a rat cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FDPDPF in its extracellular region and the
CC sequence YWFM in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.
XX
SQ Sequence 200 AA;

Query Match 100.0%; Score 6; DB 19; Length 200;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FDPDPF 6
| | | | |
Db 114 fdpppf 119

RESULT 8
AAB08723
AAB08723 standard; Protein; 200 AA.

XX AC AAB08723;

```

XX DE 02-JAN-2001 (first entry)
XX DE Amino acid sequence of a murine CRP1 polypeptide.
XX KW CRP1; CD28 related protein-1; B7RP1; B7 related protein-1;
XX KW T-lymphocyte activation; type I transmembrane protein; T cell activation;
XX KW T cell proliferation; T-cell mediated disorder.
XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT Peptide 1..20
XX FT Protein /note= "signal peptide"
XX FT Domain 21..200
XX FT Domain /note= "mature protein"
XX FT Domain 21..145
XX FT Domain /note= "extracellular domain"
XX FT Domain 146..163
XX FT Domain /note= "predicted transmembrane domain"
XX FT Domain 164..200
XX FT Domain /note= "intracellular domain"
XX PN WO200046240-A2.
XX PD 10-AUG-2000.
XX PF 27-JAN-2000; 2000WO-US01871.
XX PR 03-FEB-1999; 99US-0244448.
XX PR 08-MAR-1999; 99US-0264527.
XX PA (AMGE-) AMGEN INC.
XX PI Yoshinaga SK;
XX DR WPI; 2000-543476/49.
XX DR N-PSDB; AAA64554.
XX PT Novel nucleic acids encoding the proteins CRP-1 and B7RP1 are useful
XX PT in the treatment, prevention and diagnosis of T cell mediated disorders
XX PS Claim 11; Fig 1A; 174pp; English.
XX CC The present sequence represents a CRP1 (CD28 related protein-1)
XX CC polypeptide. The specification also describes a B7RP1 (B7 related
XX CC protein-1) polypeptide. The polypeptides are involved in T-lymphocyte
XX CC activation, and represent a receptor-ligand pair. CRP1 and B7RP1 are
XX CC predicted to be a type I transmembrane protein. The nucleic acids are
XX CC useful for regulating T cell activation or proliferation in an animal.
XX CC The polypeptides are useful for treating, preventing ameliorating or
XX CC diagnosing a T-cell mediated disorder in an animal. They can also be
XX CC used to identify test molecules that bind to the polypeptides.
XX SQ Sequence 200 AA;
XX Query Match 100.0%; Score 6; DB 21; Length 200;
XX Best Local Similarity 100.0%; Pred. No. 5.1;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 FDPDPF 6
XX Db 114 fdpppf 119
XX RESULT 9
XX ID AAY92212
XX AC AAY92212 standard; Protein; 200 AA.
XX AC AAY92212;
XX XX
XX DT 10-AUG-2000 (first entry)

```

```

XX DE Murine Th2-specific polypeptide, ml228.
XX DE ml228; Th2-specific; T helper cell; anti-inflammatory; antiarthritic;
XX KW CD28; antidiabetic; immunosuppressive; neuroprotective; dermatological;
XX KW antipsoriatic; antiasthmatic; anti-allergic; anti-viral; ophthalmological;
XX KW CTLA-4; nephrotropic; anti-HIV; antibacterial.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
XX FT Peptide 114..119
XX FT Protein /label= "conserved ppp motif"
XX FT Protein /note= "common in CD28 and CTLA-4"
XX FT Peptide 181..184
XX FT Protein /label= "YXXM motif"
XX FT Protein /note= "common in CD28 and CTLA-4; necessary for
XX FT CD28-mediated phosphatidylinositol 3-kinase
XX FT activity"
XX PN WO200019988-A1.
XX PD 13-APR-2000.
XX PF 06-OCT-1999; 99WO-US23156.
XX PR 07-OCT-1998; 98US-0168229.
XX PR 26-FEB-1999; 99US-0258670.
XX PR 06-OCT-1999; 99US-0413136.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Lehar S, Manning S, Coyle AJ, Gutierrez-ramos J;
XX DR WPI; 2000-303619/26.
XX DR N-PSDB; AAA09054.
XX PT T helper (Th) 2 nucleic acids and encoded proteins, useful for the
XX PT diagnosis and treatment of immune and respiratory disorders such as
XX PT Crohn's diseases, arthritis, insulin dependent diabetes and
XX PT autoimmunity
XX PS Claim 16; Page 130-131; 159pp; English.
XX CC This Th2-specific polypeptide is encoded by a murine orthologue ml288.
XX CC The protein shares homology with both human and murine CD28 and CTLA-4.
XX CC A novel method for modulating a Th2 response, an immune response, or
XX CC suppressing airway inflammation or hyperresponsiveness in a mammal
XX CC comprises administering a Th2-specific polypeptide of the invention, an
XX CC antibody to such a polypeptide or allelic variants of the genes. The
XX CC novel DNA and polypeptide sequences are useful for treatment and
XX CC diagnosis of Th cell and Th cell-like disorders such as Crohn's diseases,
XX CC arthritis, insulin dependent diabetes, autoimmunity, multiple sclerosis,
XX CC dermatitis, psoriasis, graft rejection, graft versus host diseases,
XX CC sarcoidosis, asthma, allergies, conjunctivitis, glomerular nephritis,
XX CC viral infections (including human immunodeficiency virus (HIV)),
XX CC bacterial infections, bronchitis, cystic fibrosis, diptheria,
XX CC emphysema, pneumonia, and Legionnaires disease.
XX SQ Sequence 200 AA;
XX Query Match 100.0%; Score 6; DB 21; Length 200;
XX Best Local Similarity 100.0%; Pred. No. 5.1;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 FDPDPF 6
XX Db 114 fdpppf 119
XX RESULT 10
XX AAW71875

```

ID - AAW71875 standard; Protein; 216 AA.
XX
AC AAW71875;
XX
DT 11-DEC-1998 (first entry)
XX
DE Rat cell surface protein #2.
XX
DE Rat; cell surface protein; thymocyte; lymphocyte; cell adhesion;
KW signal transduction; autoimmune disorder; allergy; diagnosis;
KW mitogen-stimulated.
XX
OS Rattus sp.
XX
PN WO9838216-A1.
XX
PD 03-SEP-1998.
XX
PD 27-FEB-1998; 98WO-JP00837.
PF 26-FEB-1998; 98JP-0062217.
XX 27-FEB-1997; 97JP-0062290.
XX (NIBS) JAPAN TOBACCO INC.
PA Tamatani T, Tezuka K;
XX
PI WPI: 1998-481144/A1.
DR N-PSDB; AAV61358.
XX
XX Cell surface molecule expressed in thymocytes and lymphocytes and -
PT mediating signal transduction and cell adhesion, and antibodies to
PT it useful in treatment of auto-immune and allergic disorders.
XX
PS Claim 9; Page 112-115; 149pp; Japanese.
XX
CC The present sequence represents a rat cell surface protein which is
CC expressed by thymocytes and by mitogen-stimulated lymphocytes. The cell
CC surface protein induces adhesion of mitogen-stimulated lymphocytes to
CC antibodies recognising the cell surface protein. These antibodies also
CC produce an increase in peripheral blood lymphocytes in the presence of
CC an antibody recognising CD3 antigen. The cell surface protein contains
CC the amino acid sequence FDPDPF in its extracellular region and the
CC sequence YMFV in its intracellular region. The cell surface protein can
CC be used in the prevention and treatment of autoimmune and allergic
CC diseases, and in the diagnosis and investigation of such disorders.
XX
XX Sequence 216 AA;
Query Match 100.0%; Score 6; DB 19; Length 216;
Best Local Similarity 100.0%; Pred. No. 5.4; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;
QY 1 FDPDPF 6
DB 114 fdpppf 119
RESULT 11
AAP80874
ID AAP80874 standard; protein; 10 AA.
XX
AC AAP80874;
XX
DT 06-NOV-1990 (first entry)
XX
DE Hook region #5 contg cleavage site and hinge region.
XX
OS Synthetic.
XX
KW hook region; peptide cleavage site; hinge region.
XX

FH Key Location/Qualifiers
FT Cleavage-site 1..1
FT /label=potential cleavage site
FT /note="linked to leader encoding beta-gal"
FT Cleavage-site 10..10
FT /label=potential cleavage site
FT /note="linked to calcitonin"
FT Region 2..9
FT /label=hinge
FT /note="promotes cleavage"
XX
PN WO8807085-A.
XX
PD 22-SEP-1988.
XX
PD 04-MAR-1988; 88WO-US00717.
XX
PR 20-MAR-1987; 87US-0028484.
XX
PA (CREA-) CREATIVE BIOMOLECUL.
XX
PI Huston JS, Cohen CM, Keck PC, Rueger DC, Charette M, Crea R;
PI Oppermann H, Ridge RJ;
XX
DR WPI: 1988-285539/40.
XX
XX Prodn of recombinant polypeptides -
PT using fused polypeptide having leader sequence, hinge region,
PT cleavage site and target polypeptide sequence.
XX
PS Disclosure; 5pp; English.
XX
CC A fusion polypeptide was synthesised in which calcitonin
CC was linked to this hook region via Arg at posn 4. The leader
CC sequence to which the hook is joined encodes beta-galactosidase.
CC Calcitonin is obtained following cleavage by Arg-C or V-8
CC protease. The hook polypeptide also contains a hinge
CC region (the hinge is a short Cys-free region which promotes cleavage).
CC See also AAP80869-P80873 and AAP80875-P80879.
XX
SQ Sequence 10 AA;
Query Match 83.3%; Score 5; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FDPDP 5
DB 2 fdppp 6
RESULT 12
AAP80875
ID AAP80875 standard; protein; 13 AA.
XX
AC AAP80875;
XX
DT 06-NOV-1990 (first entry)
XX
DE Hook region #6 contg cleavage sites and hinge region.
XX
OS Synthetic.
XX
KW hook region; peptide cleavage site; hinge region.
XX
FH Key Location/Qualifiers
FT Cleavage-site 1..1
FT /label=potential cleavage site
FT /note="linked to beta-gal leader"
FT Cleavage-site 13..13
FT /label=potential cleavage site
FT /note="linked to human calcitonin"

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:45:16 ; Search time 25.43 Seconds
(without alignments)
4.858 Million cell updates/sec

Title: US-09-383-551B-21

Perfect score: 6

Sequence: 1 FDPFPF 6

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents,AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5b_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	83.3	126	2	US-08-341-843B-36
2	5	83.3	126	2	US-08-427-497B-41
3	5	83.3	140	3	US-08-850-961-12
4	5	83.3	141	3	US-08-850-961-10
5	5	83.3	409	2	US-08-809-740A-5
6	4	66.7	7	4	US-09-024-975-5
7	4	66.7	7	4	US-09-147-933-35
8	4	66.7	8	2	US-09-016-366A-38
9	4	66.7	9	4	US-08-185-883-7
10	4	66.7	9	4	US-09-203-921-7
11	4	66.7	16	4	US-09-024-975-3
12	4	66.7	16	4	US-08-602-999A-181
13	4	66.7	16	4	US-08-602-999A-197
14	4	66.7	16	4	US-08-602-999A-199
15	4	66.7	16	4	US-08-602-999A-440
16	4	66.7	18	4	US-08-602-999A-409
17	4	66.7	19	4	US-08-781-420-4
18	4	66.7	20	4	US-08-602-999A-152
19	4	66.7	20	4	US-08-602-999A-154
20	4	66.7	24	2	US-08-769-745-21
21	4	66.7	26	2	US-08-419-066-2
22	4	66.7	26	4	US-09-024-975-2
23	4	66.7	28	2	US-08-598-873-29
24	4	66.7	28	4	US-08-605-430-29
25	4	66.7	32	2	US-08-637-759B-343
26	4	66.7	32	3	US-08-871-355A-343
27	4	66.7	36	1	US-08-487-359-9

28	4	66.7	36	1	US-08-222-798A-9	Sequence 9, Appli
29	4	66.7	39	1	US-08-162-052-1	Sequence 1, Appli
30	4	66.7	39	1	US-08-310-722-1	Sequence 1, Appli
31	4	66.7	39	2	US-08-419-066-1	Sequence 1, Appli
32	4	66.7	39	2	US-08-728-333-1	Sequence 1, Appli
33	4	66.7	39	4	US-09-024-975-1	Sequence 1, Appli
34	4	66.7	39	5	PCT-US95-12080-1	Sequence 1, Appli
35	4	66.7	51	1	US-08-188-228-38	Sequence 38, Appli
36	4	66.7	51	1	US-08-188-228-40	Sequence 40, Appli
37	4	66.7	51	1	US-08-332-643-38	Sequence 38, Appli
38	4	66.7	51	1	US-08-332-643-40	Sequence 40, Appli
39	4	66.7	51	1	US-08-332-638-38	Sequence 38, Appli
40	4	66.7	51	1	US-08-332-638-40	Sequence 40, Appli
41	4	66.7	56	1	US-07-972-481-4	Sequence 4, Appli
42	4	66.7	78	1	US-08-487-359-5	Sequence 5, Appli
43	4	66.7	78	1	US-08-487-359-7	Sequence 7, Appli
44	4	66.7	78	1	US-08-222-798A-5	Sequence 5, Appli
45	4	66.7	78	1	US-08-222-798A-7	Sequence 7, Appli
46	4	66.7	79	1	US-08-487-359-1	Sequence 1, Appli
47	4	66.7	79	1	US-08-487-359-2	Sequence 2, Appli
48	4	66.7	79	1	US-08-487-359-3	Sequence 3, Appli
49	4	66.7	79	1	US-08-487-359-4	Sequence 4, Appli
50	4	66.7	79	1	US-08-487-359-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-341-843B-36
; Sequence 36, Application US/08341843B
; Patent No. 5872225
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the
; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
; Patent No. 5872225
; TITLE OF INVENTION: the Nucleotide Sequence
; TITLE OF INVENTION: Characterized Thereby
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fay, Sharpe, Beall, Fagan,
; ADDRESSEE: Minnich & McKee
; STREET: 1100 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2518
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; MEDIUM TYPE: storable
; COMPUTER: Compaq Prolinea 5100e
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/341,843B
; FILING DATE: No. 5872225ember 18, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/904,991
; FILING DATE: June 26, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Minnich, Richard J.
; REGISTRATION NUMBER: 24,175
; REFERENCE/DOCKET NUMBER: CWR 2 149-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 861-5582
; TELEFAX: (216) 241-1666
; TELEX: (216) 980162
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:46:04 ; Search time 32.53 Seconds
(without alignments)
9.367 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YNFM 4

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database: PIR_68:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	118	2 S72690	probable membrane
2	4	100.0	199	2 S78540	inducible T-cell c
3	4	100.0	200	2 JC7397	activation-inducib
4	4	100.0	216	2 JC7396	activation-inducib
5	4	100.0	219	2 F71710	hypothetical prote
6	4	100.0	256	2 T25707	hypothetical prote
7	4	100.0	269	2 T24112	hypothetical prote
8	4	100.0	289	2 A70751	hypothetical prote
9	4	100.0	314	2 T45077	ornithine carbamoy
10	4	100.0	317	2 G71119	probable ornithine
11	4	100.0	317	2 G75041	ornithine carbamoy
12	4	100.0	334	2 G84123	iron (III) dicitra
13	4	100.0	335	2 T23350	hypothetical prote
14	4	100.0	348	2 T11301	NADH dehydrogenase
15	4	100.0	351	2 S29152	rhodopsin - chicke
16	4	100.0	355	2 A75050	hypothetical prote
17	4	100.0	358	2 F71110	hypothetical prote
18	4	100.0	426	2 B96519	protein T2E6.2 [i
19	4	100.0	429	2 A86486	protein F28J9.8 [i
20	4	100.0	431	2 F71600	variant-specific s
21	4	100.0	455	2 H69837	conserved hypothet
22	4	100.0	516	2 B64902	probable oligopept
23	4	100.0	516	2 F85727	probable hemin-bin
24	4	100.0	645	2 A75390	NADH dehydrogenase
25	4	100.0	653	2 B71640	NADH dehydrogenase
26	4	100.0	788	2 T15130	integrin beta-3 su
27	4	100.0	792	2 G84830	probable potassium
28	4	100.0	880	2 T38083	potassium transpor
29	4	100.0	908	2 H86537	oxoglutarate dehyd

30	4	100.0	908	2 G72085	2-oxoglutarate deh
31	4	100.0	1078	2 E75407	isoleucyl-tRNA syn
32	4	100.0	1442	2 S57160	sulfite reductase
33	4	100.0	1766	2 S03701	141K protein - pea
34	4	100.0	1926	2 JC4842	DNA-binding nuclea
35	4	100.0	2207	2 T24629	glutamate synthase
36	4	100.0	2207	2 T23433	hypothetical prote
37	4	100.0	3672	2 T23433	probable laminin a
38	3	75.0	13	2 PN0122	OLL protein - vacc
39	3	75.0	13	2 S29488	GRP-binding protei
40	3	75.0	29	2 A32414	bothrolysin (EC 3.
41	3	75.0	30	2 S34765	4-hydroxybutyryl-C
42	3	75.0	37	2 H82304	hypothetical prote
43	3	75.0	38	2 T14226	NADH dehydrogenase
44	3	75.0	40	2 S29489	GRP-binding protei
45	3	75.0	40	2 D81036	hypothetical prote
46	3	75.0	41	2 G69484	hypothetical prote
47	3	75.0	45	2 F70138	hypothetical prote
48	3	75.0	46	2 S14720	beta-lactoglobulin
49	3	75.0	46	2 A23952	hypothetical prote
50	3	75.0	49	2 A42766	jararafibrase II (

ALIGNMENTS

RESULT 1

S72690

probable membrane protein Q0270 - yeast (Saccharomyces cerevisiae) mitochondrion
C:Species: mitochondrion Saccharomyces cerevisiae

C:Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2001

C:Accession: S72690

R:de Zamaroczy, M.; Bernardi, G.

Gene 47, 155-177, 1986

A:Title: The primary structure of the mitochondrial genome of Saccharomyces cerevisia

A:Reference number: S72681; MUID:87163488

A:Accession: S72690

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-118 <DEZ>

A:Note: this is not the original publication

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC2

C:Keywords: mitochondrion; transmembrane protein

F:17-33/Domain: transmembrane #status predicted <TM1>

F:49-65/Domain: transmembrane #status predicted <TM2>

F:73-89/Domain: transmembrane #status predicted <TM3>

Query Match 100.0%; Score 4; DB 2; Length 118;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YNFM 4

||||

Db 93 YNFM 96

RESULT 2

S78540

Inducible T-cell co-stimulator ICOS precursor - human

C:Species: Homo sapiens (man)

C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 07-May-1999

C:Accession: S78540; S78748; S78749

R:Krocsek, R.

submitted to the Protein Sequence Database, June 1998

A:Reference number: S78540

A:Accession: S78540

A:Molecule type: mRNA

A:Residues: 1-199 <KRO>

A:Experimental source: cell line MOLT-4V

R:Hutloff, A.; Dittlich, A.M.; Beier, K.C.; Eljaschewitsch, B.; Kraft, R.; Anagnostop

Nature 397, 263-266, 1999
A:Title: ICOS is an inducible T-cell co-stimulator structurally and functionally related
A:Reference number: S78748; MUID:99127892
A:Accession: S78748
A:Molecule type: mRNA
A:Residues: 1-199 <HUT1>
A:Experimental source: cell line MOLT-4V
A:Accession: S78749
A:Molecule type: protein
A:Residues: 'X',193-198 <HUT2>
A:Experimental source: cell line MOLT-4V
C:Complex: homodimer
C:Superfamily: Immunoglobulin homology
C:Keywords: dimer; glycoprotein; T-cell; transmembrane protein
F:1-19/Domain: (or 1-20) signal sequence #status predicted <SIG>
F:20-199/Product: (or 21-199) inducible costimulator ICOS #status predicted <MAT>
F:21-138/Domain: extracellular #status predicted <EXT>
F:26-132/Domain: immunoglobulin homology <IMM>
F:139-164/Domain: transmembrane #status predicted <INT>
F:165-199/Domain: intracellular #status predicted <INT>
F:189,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 4; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 180 YMPM 183

RESULT 3
JC7397
activation-inducible lymphocyte immuno-mediatory molecule-2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7397; PC7100
R:Tezuka, K.; Tsuji, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kanada, Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat AILIM/ICOS, a novel T-cell costimulator
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7397
A:Molecule type: mRNA
A:Residues: 1-200 <TEZ>
A:Cross-references: DDBJ:AB023134
A:Accession: PC7100
A:Molecule type: protein
A:Residues: 21-40 <PE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
A:Gene: ailim-2
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 4; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 181 YMPM 184

RESULT 4
JC7396
activation-inducible lymphocyte immuno-mediatory molecule-1 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7396; PC7099
R:Tezuka, K.; Tsuji, T.; Hirano, D.; Tamatani, T.; Sakamaki, K.; Kobayashi, Y.; Kanada,

Biochem. Biophys. Res. Commun. 276, 335-345, 2000
A:Title: Identification and characterization of rat AILIM/ICOS, a novel T-cell costim
A:Reference number: JC7396
A:Contents: Spleen
A:Accession: JC7396
A:Molecule type: mRNA
A:Residues: 1-216 <TEZ>
A:Cross-references: DDBJ:AB023133
A:Accession: PC7099
A:Molecule type: protein
A:Residues: 21-40 <PE2>
C:Comment: This protein is an inducible cell surface glycoprotein, and a type I transmembrane protein, and plays an important role for the maturation or selection of T cells in thymus.
C:Genetics:
A:Gene: ailim-1
C:Keywords: glycoprotein; T-cell; transmembrane protein

Query Match 100.0%; Score 4; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 181 YMPM 184

RESULT 5
F71710
hypothetical protein RP030 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: F71710
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alismark Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: F71710
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-219 <AND>
A:Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAAL4501.1; PID:e134
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: RP030
C:Superfamily: Rickettsia prowazekii hypothetical protein RP030

Query Match 100.0%; Score 4; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
|||||
Db 153 YMPM 156

RESULT 6
T25707
hypothetical protein F19F10.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25707
R:Kellen, J.; Wamsley, P.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid F19F10.
A:Reference number: Z20073
A:Accession: T25707
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-256 <REL>
A:Cross-references: EMBL:U97005; PIDN:AAB52280.1; GSPDB:GN00023; CESP:F19F10.2
A:Experimental source: strain Bristol N2; clone F19F10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:49 ; Search time 17.42 Seconds
(without alignments)
7.866 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YMFM 4

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	289	1 Y097_MYCTU	Q10893 mycobacteri
2	4	100.0	315	1 OTC_PYRFU	O51742 pyrococcus
3	4	100.0	317	1 OTC_PYRAB	O93656 pyrococcus
4	4	100.0	317	1 OTC_PYRHO	O58457 pyrococcus
5	4	100.0	345	1 GDFB_RAT	Q92217 rattus norv
6	4	100.0	348	1 NU2M_SCYCA	O21409 scyliorhinu
7	4	100.0	348	1 OPSD_PIG	O18766 sus scrofa
8	4	100.0	351	1 OPSD_CHICK	P22328 gallus gall
9	4	100.0	352	1 OPSD_ALLMI	P52202 alligator m
10	4	100.0	405	1 GDFB_MOUSE	Q921w4 mus musculu
11	4	100.0	407	1 GDFB_HUMAN	O95390 homo sapien
12	4	100.0	511	1 NU4M_HANWI	P48917 hansenua w
13	4	100.0	516	1 YDGS_ECOLI	P76128 escherichia
14	4	100.0	880	1 TRK2_SCHPO	Q10065 schizosacch
15	4	100.0	1442	1 YJ9F_YEAST	P47169 saccharomyc
16	4	100.0	3672	1 LMJ2_CAEEL	Q21313 caenorhabdi
17	3	75.0	9	1 FLA2_TREHY	P80159 treponema h
18	3	75.0	29	1 HRJ_BOTJA	P20416 bothrops ja
19	3	75.0	30	1 HDVD_CLOAB	P55792 clostridium
20	3	75.0	41	1 S482_HVSGI	P56854 hysteroerat
21	3	75.0	45	1 Y310_BORBU	O51290 borrelia bu
22	3	75.0	55	1 ATP8_STRCA	O21401 struthio ca
23	3	75.0	56	1 NUGM_ARTSA	P19048 artemia sal
24	3	75.0	60	1 CX10_NAJHA	P01453 naja haje a
25	3	75.0	60	1 CX1_NAJHA	P01455 naja haje a
26	3	75.0	60	1 CX1_NAJNI	P01456 naja nivea
27	3	75.0	60	1 CX1_NAJOX	P01451 naja oxiana
28	3	75.0	60	1 CX2_NAJKA	P01445 naja naja k
29	3	75.0	60	1 CX2_NAJOX	P01441 naja oxiana
30	3	75.0	60	1 CX3_NAJKA	P01446 naja naja k
31	3	75.0	60	1 CX3_NAJNA	P24780 naja naja (
32	3	75.0	60	1 CX3T_NAJAT	P07525 naja atra (
33	3	75.0	60	1 CX3S_NAJKA	P24779 naja naja k

34	3	75.0	60	1 CX6_NAJAT	P80245 naja atra (
35	3	75.0	60	1 CX9_NAJHA	P01454 naja haje a
36	3	75.0	61	1 RL21_PIG	P49666 sus scrofa
37	3	75.0	68	1 YHVH_SALTY	P40819 salmonella
38	3	75.0	69	1 YVCA_VACCC	P21119 vaccinia vi
39	3	75.0	70	1 ATP_L_BACSU	P37815 bacillus su
40	3	75.0	72	1 MTRG_METMA	P80656 methanosarc
41	3	75.0	77	1 THGL_ARATH	O39182 arabidopsis
42	3	75.0	78	1 Y8K2_SSVI	P20205 sulfolobus
43	3	75.0	81	1 CX1B_NAJAT	O98956 naja atra (
44	3	75.0	81	1 CX1C_NAJAT	O98957 naja atra (
45	3	75.0	81	1 CX1D_NAJAT	O98958 naja atra (
46	3	75.0	81	1 CX1_NAJAT	P01449 naja atra (
47	3	75.0	81	1 CX1_NAJSP	O02454 naja sputat
48	3	75.0	81	1 CX2_NAJAT	P01442 naja atra (
49	3	75.0	81	1 CX3A_NAJAT	O98959 naja atra (
50	3	75.0	81	1 CX3B_NAJAT	O98960 naja atra (

ALIGNMENTS

RESULT 1					
Y097_MYCTU					
ID Y097_MYCTU	STANDARD;	PRT;	289 AA.		
AC Q10893;					
DT 01-OCT-1996 (Rel. 34, Created)					
DT 01-OCT-1996 (Rel. 34, Last sequence update)					
DT 30-MAY-2000 (Rel. 39, Last annotation update)					
DE HYPOTHETICAL 32.6 KDA PROTEIN RV0097.					
GN RV0097 OR WTCY251.16					
OS Mycobacterium tuberculosis.					
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;					
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.					
OX NCBI_TaxID=1773;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=H37RV;					
RX MEDLINE=98295987; PubMed=9634230;					
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,					
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,					
RA Radcock K., Basham D., Brown D., Chillingworth T., Connor R.,					
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,					
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,					
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,					
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,					
RA Taylor K., Whitehead S., Barrell B.G.;					
RT "Deciphering the biology of Mycobacterium tuberculosis from the					
RT complete genome sequence."					
RL Nature 393:537-544(1998).					
CC -!- SIMILARITY: SOME, TO ALCALIGENES EUTROPHUS 2.4-					
CC DICHLOPHENOXACETATE MONOOXYGENASE (TFDA).					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC or send an email to license@isb-sib.ch).					
CC -----					
CC EMBL; 274410; CAA989933.1; -					
DR Tuberculinist; RV0097; -					
DR KW Hypothetical protein; Oxidoreductase.					
SQ SEQUENCE 289 AA; 32641 MW; FD38D0283E0E9F95 CRC64;					

Query Match 100.0%; Score 4; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMFM 4
|||||

Db 100 YMEM 103

```
RESULT 2
ID OTC_PYRFU STANDARD; PRT; 315 AA.
AC Q51742;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ORNITHINE CARBAMOYLTRANSFERASE (EC 2.1.3.3) (OTCASE).
GN ARGF.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=DSM 3638 / VC1;
RC MEDLINE=97433289; PubMed=9288929;
RC Roovers M., Hetcke C., Legrain C., Thomm M., Glansdorff N.;
RC "Isolation of the gene encoding Pyrococcus furiosus ornithine
RC carbamoyltransferase and study of its expression profile in vivo and
RC in vitro.";
RC Eur. J. Biochem. 247:1038-1045(1997).
RN [2]
RA Roovers M.;
RA REVISION TO 207.
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=98169451; PubMed=95011170;
RA Villaret V., Clantin B., Tricot C., Legrain C., Roovers M., Stalon V.,
RA Glansdorff N., van Beeumen J.;
RA "The crystal structure of Pyrococcus furiosus ornithine
RA carbamoyltransferase reveals a key role for oligomerization in enzyme
RA stability at extremely high temperatures.";
RA Proc. Natl. Acad. Sci. U.S.A. 95:2801-2806(1998).
CC -1- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE -> CITRULLINE
CC + ORTHOPHOSPHATE.
CC -1- PATHWAY: ARGinine BIOSYNTHESIS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X99225; CAA67609.1; -
DR EMBL: Y12727; CAA73260.1; -
DR PDB: 1A1S; 17-JUN-98.
DR InterPro: IPR002029; -
DR InterPro: IPR002082; -
DR InterPro: IPR002292; -
DR Pfam: PF00185; OTCase; 1.
DR PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
DR TRANSFERASE; Arginine biosynthesis.
DR TRANSFERASE; Arginine biosynthesis.
FT CONFLICT 207 207 MISSING (IN CAA67609).
FT CONFLICT 315 AA; 35280 MW; 62804E1F2A89E81 CRC64;
SQ SEQUENCE 315 AA; 35280 MW; 62804E1F2A89E81 CRC64;
```

Query Match 100.0%; Score 4; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMEM 4
||||

Db 265 YMEM 268

```
RESULT 3
ID OTC_PYRAB STANDARD; PRT; 317 AA.
AC Q93656;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ORNITHINE CARBAMOYLTRANSFERASE (EC 2.1.3.3) (OTCASE).
GN ARGF OR PAB1502.
OS Pyrococcus abyssii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=GE5;
RC Cunin R., Seumois G., Purcareau C., Van Vliet F., Legrain C.;
RC "Ornithine carbamoyltransferase from Pyrococcus abyssii.";
RC Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RA "Pyrococcus abyssii genome sequence: insights into archaeal chromosome
RA structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: CARBAMOYL PHOSPHATE + ORNITHINE -> CITRULLINE
CC + ORTHOPHOSPHATE.
CC -1- PATHWAY: ARGinine BIOSYNTHESIS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF083209; AAD09004.1; -
DR EMBL: AJ248287; CAB50228.1; -
DR HSSP: Q51742; 1A1S.
DR InterPro: IPR002029; -
DR InterPro: IPR002082; -
DR InterPro: IPR002292; -
DR Pfam: PF00185; OTCase; 1.
DR PRINTS: PR00100; AOTCase.
DR PRINTS: PR00101; ATCase.
DR PRINTS: PR00102; OTCase.
DR PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
DR TRANSFERASE; Arginine biosynthesis.
DR TRANSFERASE; Arginine biosynthesis.
FT CONFLICT 63 63 S -> R (IN REF. 1).
FT CONFLICT 75 75 A -> G (IN REF. 1).
FT CONFLICT 159 159 V -> I (IN REF. 1).
FT CONFLICT 201 202 EQ -> DE (IN REF. 1).
FT CONFLICT 264 264 D -> S (IN REF. 1).
FT CONFLICT 313 317 VKTGF -> RKDGLLTF (IN REF. 1).
SQ SEQUENCE 317 AA; 35357 MW; F4686B091F76494F CRC64;
```

Query Match 100.0%; Score 4; DB 1; Length 317;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMEM 4
||||

RESULT 4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:52:21 ; Search time 47.57 Seconds

(without alignments)

11.125 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YMFH 4

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	4	100.0	85	Q36059	trachyrhach
2	4	100.0	89	Q9G123	myotis adve
3	4	100.0	100	Q47726	gazella ruf
4	4	100.0	102	Q9TEF4	podarcis at
5	4	100.0	102	Q9TEF3	podarcis at
6	4	100.0	120	Q9PUR0	petromyzon
7	4	100.0	124	Q03232	asobara tab
8	4	100.0	128	Q47597	atolocynus
9	4	100.0	131	Q47601	chrysocyon
10	4	100.0	131	Q9T9E6	aquarius el
11	4	100.0	131	Q9T9E5	aquarius el
12	4	100.0	132	Q47616	speothos ve
13	4	100.0	134	Q19030	mogera wogu
14	4	100.0	134	Q021395	scoriculus c
15	4	100.0	134	Q9MQY2	euroscaptor
16	4	100.0	134	Q9MQX9	mogera toku
17	4	100.0	134	Q9MQX7	mogera imai
18	4	100.0	134	Q9MQX5	mogera wogu
19	4	100.0	134	Q9MQX4	mogera wogu

20	4	100.0	134	8	Q9MQX3	Q9mqx3 mogera wogu
21	4	100.0	134	8	Q9MQX2	Q9mqx2 mogera wogu
22	4	100.0	134	8	Q9MQX0	Q9mqx0 mogera wogu
23	4	100.0	134	8	Q9ME45	Q9me45 mogera wogu
24	4	100.0	134	8	Q9MDM2	Q9mdm2 mogera imai
25	4	100.0	134	8	Q9MDG3	Q9mdg3 mogera imai
26	4	100.0	134	8	Q9MDF1	Q9mdf1 mogera wogu
27	4	100.0	134	8	Q9MDF0	Q9mdf0 mogera wogu
28	4	100.0	155	8	Q79070	Q79070 pauesia sp.
29	4	100.0	155	8	Q79071	Q79071 pauesia sil
30	4	100.0	155	8	Q79081	Q79081 trioxys fal
31	4	100.0	160	14	P89293	P89293 xestia c-ni
32	4	100.0	199	4	Q9Y6W8	Q9y6w8 homo sapien
33	4	100.0	200	11	Q9R1F7	Q9r1f7 ratus norv
34	4	100.0	200	11	Q9WVS0	Q9wvs0 mus musculu
35	4	100.0	200	11	Q9JL17	Q9jll17 mus musculu
36	4	100.0	203	5	O16010	O16010 drosophila
37	4	100.0	212	14	Q9W105	Q9w105 bovine herp
38	4	100.0	216	11	Q9WVR9	Q9wvr9 ratus norv
39	4	100.0	219	2	Q9ZEB5	Q9zeb5 rickettsia
40	4	100.0	247	10	O24659	O24659 oryza sativ
41	4	100.0	256	5	O01518	O01518 caenorhabdi
42	4	100.0	269	5	Q9XVK3	Q9xvk3 caenorhabdi
43	4	100.0	305	2	O50372	O50372 methylophil
44	4	100.0	320	13	Q9I971	Q9i971 serinus can
45	4	100.0	322	13	O57448	O57448 anas platyr
46	4	100.0	325	14	Q9PYL6	Q9pyl6 xestia c-ni
47	4	100.0	334	2	Q9K6D8	Q9k6d8 bacillus ha
48	4	100.0	335	5	O45660	O45660 caenorhabdi
49	4	100.0	347	8	O21177	O21177 heteroceph
50	4	100.0	351	13	O57450	O57450 melopsittac

ALIGNMENTS

RESULT 1	
ID Q36059	PRELIMINARY; PRT; 85 AA.
AC Q36059;	
DT 01-NOV-1996 (TrEMBLrel. 01, Created)	
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)	
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)	
DE CYTOCHROME B (FRAGMENT).	
GN PETB OR PETD.	
OS Trachyrhachis kiowa (Kiowa grasshopper).	
OG Mitochondrion.	
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;	
OC Acridomorpha; Acridoidea; Acrididae; Trachyrhachis.	
OX NCBI_TaxID=37263;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=LAST MOUNTAIN LAKE, SASKATCHEWAN;	
RC TISSUE=WHOLE ORGANISM MINUS GUT;	
RA Chapco W., Martel R.K.B.;	
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.	
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE	
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A	
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL	
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).	
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY	
CC BOUND TO THE PROTEIN (BY SIMILARITY).	
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,	
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).	
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.	
DR EMBL; U18255; AAA61676.1; -.	
DR InterPro; IPR000179; -.	
DR pfam; PF00033; cytochrome_b_n: 1.	
KW Electron transport; Heme; Mitochondrion; Respiratory chain;	
KW Transmembrane.	
FT NON_TER 1 1	
FT NON_TER 85 85	

SQ • SEQUENCE 85 AA; 9584 MW; DC97A7512F8590FC CRC64;

Query Match 100.0%; Score 4; DB 8; Length 85;

• Best Local Similarity 100.0%; Pred. No. 78; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
||||

Db 30 YMF 33

RESULT 2

ID Q9G123 PRELIMINARY; PRT; 89 AA.

AC Q9G123;

DT 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)

DE CYTOCHROME B (FRAGMENT).

QYTB

OS Myotis adersus (large-footed bat).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Vespertilionidae;

OC Myotis.

OX NCBI_TaxID=59461;

RN [1]

RP SEQUENCE FROM N.A.

RA Cooper S.J.B., Day P.R., Reardon T.B., Schulz M.;

RT "Assessment of species boundaries in Australian Myotis (Chiroptera:

RT Vespertilionidae) using mitochondrial DNA.";

RL J. Mammal. 0:0-0(2000);

DR EMBL; AY007529; AAG27054.1; -;

DR EMBL; AY007528; AAG27053.1; -;

KW Mitochondrion.

FT NON_TER 1

FT NON_TER 89

SQ SEQUENCE 89 AA; 10127 MW; 48D12EC6FB0AFC0 CRC64;

Query Match

Best Local Similarity 100.0%; Score 4; DB 8; Length 89;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
||||

69 YMF 72

RESULT 3

047726

ID 047726 PRELIMINARY; PRT; 100 AA.

AC 047726;

DT 01-JUN-1998 (TRENBLrel. 06, Created)

DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)

DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)

DE CYTOCHROME B (FRAGMENT).

GN PETB OR PETD OR CYT B.

OS Gazella rufifrons (Red-fronted gazelle).

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Antilopinae; Gazella.

OX NCBI_TaxID=69304;

RN [1]

RP SEQUENCE FROM N.A.

RA Rebholz W.E.R., Harley E.H.;

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC

CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C → Q + 2
CC FERROCYTOCHROME C.
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.
DR EMBL; AF030606; AAB93577.1; -;
DR InterPro; IPR000179; -;
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
FT NON_TER 1
FT NON_TER 100
SQ SEQUENCE 100 AA; 11332 MW; 3AF07394E6160BB1 CRC64;

Query Match 100.0%; Score 4; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 87; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
||||

Db 91 YMF 94

RESULT 4

ID Q9TEF4 PRELIMINARY; PRT; 102 AA.

AC Q9TEF4;

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)

DE CYTOCHROME B (FRAGMENT).

GN PETB OR PETD.

OS Podarcis atrata.

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosauria; Squamata; Scieroglossa; Scincomorpha; Lacertoldea;

OC Lacertidae; Podarcis.

OX NCBI_TaxID=73148;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FORAD3;

RA Castilia A.M., Fernandez-Pedrosa V., Bacheljau T., Gonzalez A.,

RA Latorre A., Moya A.;

RT "Conservation genetics of insular Podarcis lizards using partial

RT cytochrome b sequences.";

RL Mol. Ecol. 7:1407-1411(1998).

CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A

CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL

CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY

CC BOUND TO THE PROTEIN (BY SIMILARITY).

CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,

CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.

DR EMBL; AJ004990; CAA06264.1; -;

DR InterPro; IPR000179; -;

DR Pfam; PF00033; cytochrome_b_n; 1.

DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain;

KW Transmembrane.

FT NON_TER 1

FT NON_TER 102

SQ SEQUENCE 102 AA; 11527 MW; D44334464C6CA15B CRC64;

Query Match 100.0%; Score 4; DB 8; Length 102;
Best Local Similarity 100.0%; Pred. No. 88; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:44:40 ; Search time 45.63 Seconds
(without alignments)
5.314 Million cell updates/sec

Title: US-09-383-551B-22

Perfect score: 4

Sequence: 1 YWFM 4

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database :

A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	100.0	126	16	AA66147
2	4	100.0	126	17	AA88554
3	4	100.0	126	18	AA23589
4	4	100.0	126	19	AA65459
5	4	100.0	126	19	AA40816
6	4	100.0	126	20	AA06096
7	4	100.0	126	20	AA06099
8	4	100.0	126	21	AA77565
9	4	100.0	128	21	AA13329
10	4	100.0	198	21	AA92213
11	4	100.0	199	19	AA75956

12	4	100.0	199	19	AA75957	Human cell surface
13	4	100.0	199	20	AA708026	Human activated T-
14	4	100.0	199	21	AA808731	Amino acid sequenc
15	4	100.0	200	19	AA75958	Mouse cell surface
16	4	100.0	200	19	AA75958	Rat cell surface p
17	4	100.0	200	21	AA808723	Amino acid sequenc
18	4	100.0	200	21	AA92212	Murine Th2-specifi
19	4	100.0	216	19	AA75957	Rat cell surface p
20	4	100.0	273	21	AA646768	Arabidopsis thalia
21	4	100.0	274	21	AA646768	Arabidopsis thalia
22	4	100.0	293	21	AA646767	Arabidopsis thalia
23	4	100.0	294	21	AA646767	Arabidopsis thalia
24	4	100.0	336	21	AA646766	Arabidopsis thalia
25	4	100.0	337	21	AA646766	Arabidopsis thalia
26	4	100.0	362	16	AA66149	Partial propeptide
27	4	100.0	362	18	AA23590	Human bone morphog
28	4	100.0	362	19	AA40817	Human bone morphog
29	4	100.0	362	20	AA06101	Human bone morphog
30	4	100.0	362	20	AA06098	Human bone morphog
31	4	100.0	407	17	AA88553	Growth differentia
32	4	100.0	407	19	AA65458	Human growth diffe
33	4	100.0	407	20	AA31195	Human GDF-11 prote
34	4	100.0	407	21	AA21088	Human GDF-11. Hom
35	4	100.0	407	21	AA92030	Human bone morphog
36	4	100.0	407	21	AA77564	Human growth diffe
37	4	100.0	407	21	AA77567	Human growth diffe
38	4	100.0	431	21	AA18350	Plasmodium falcipa
39	4	100.0	630	21	AA29366	Arabidopsis thalia
40	4	100.0	727	21	AA29365	Arabidopsis thalia
41	4	100.0	792	21	AA29364	Arabidopsis thalia
42	4	100.0	915	20	AA34977	C. pneumoniae prot
43	4	100.0	1311	19	AA52197	Precis coenia (but
44	4	100.0	1311	20	AA72971	Precis coenia patc
45	4	100.0	1311	22	AA67156	Butterfly patched
46	4	100.0	2466	20	AA05842	Banana ripening fr
47	3	75.0	4	2	AA10546	Analgesic peptide
48	3	75.0	4	5	AA40747	Dermorphine deriva
49	3	75.0	4	8	AA71581	Growth hormone rel
50	3	75.0	4	11	AA03383	Cardio-active caso

ALIGNMENTS

RESULT 1
AA66147
ID AAR66147 standard; Protein; 126 AA.
XX
AC AAR66147;
XX
DT 10-AUG-1995 (first entry)
XX
DE Partial bovine bone morphogenetic protein-11 (BMP-11).
XX
KW Bone morphogenetic protein-11; BMP-11; TGF-beta superfamily.
XX
OS Bos taurus.
FH Key Location/Qualifiers
FT Protein 18..126
FT /label= mature
XX
PN WO9426892-A.
XX
PD 24-NOV-1994.
XX
PF 12-MAY-1994; 94WO-US05288.
XX
PR 12-MAY-1993; 93US-0061464.
XX
PA (CEMY-) GENETICS INST INC.
XX
PI Celeste AJ, Wozney JM;

XX *
DR WPI: 1995-006788/01.
DR N-PSDB; AAO79444.
XX
XX New DNA encoding bone morphogenetic protein 11 - and related
PT vectors, transformed cells and polypeptide(s), including
PT heterodimers, useful e.g. in fertility control bone and tissue
PT repair, etc.
XX
XX Claim 15; Page 40-41; 57pp; English.
XX
XX A bovine genomic library (strain Bovine Activin WC) in lambda EMBL3
CC was screened under low stringency conditions with a 1081-1403 base
CC fragment of human BMP-7 DNA. Positive clones were screened with BMP-
CC 5', -6, and -7 probes under high stringency conditions and one clone
CC reactive in the first screen but not in the second was selected. The
CC hybridisation characteristics were localised to a 0.5 kb fragment.
CC The partial sequence of this clone, lambda 7r-30 (ATCCD 75439) is
CC Q79444. The 5' limit of this exon of the bovine BMP-11 gene is
CC difficult to define. Clone lambda 7r-30 contains at least one exon/
CC intron boundary. BMP-11 polypeptide exists as a dimer comprising two
CC of the mature protein AA sequences or as a heterodimer with one
CC mature sequence from BMP-11 and the other being any of BMP 1-10.
CC The predicted mol. wt. of the mature active species comprising two
CC mature protein sequences is approx. 12,000 daltons. Further active
CC species are contemplated comprising AAs 23-126. Primers C and D
CC are based on clone lambda 7r-30 (see Q79446, Q79447). Nts 375 or
CC 390-704 of Q79444 are claimed. AAs 18-126 of R66147 are claimed.
XX
XX Sequence 126 AA;

Query Match 100.0%; Score 4; DB 16; Length 126;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
| | | |
Db 66 ymf 69

RESULT 2
AAR88554
ID AAR88554 standard; Protein; 126 AA.
XX
AC AAR88554;

XX 15-APR-1996 (first entry)
DE Murine growth differentiation factor-11 (GDF-11).
XX
XX Growth differentiation factor-11; GDF-11; antibody; detection;
KW disorder; muscle; antisense; suppression; vector; liposome;
KW targeting.
XX
OS Mus musculus.
XX
PN WO9601845-A1.
XX

XX 25-JAN-1996.
XX
XX 07-JUL-1995; 95WO-US08543.
XX
XX 08-JUL-1994; 94US-0272763.
XX
XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MED.
PA Lee S, Mcperron AC;
PI
XX WPI: 1996-097589/10.
DR N-PSDB; AAT11062.
XX
XX New-Growth Differentiation Factor-11 (GDF-11) - with tissue-specific

PT expression in muscle, neural and uterine cells, for detecting cell
PT proliferation disorders
XX
XX Claim 3; Page 39-40; 67pp; English.
XX
XX Antibodies directed against the growth differentiation factor (GDF)
CC are useful for detecting cell proliferative disorders when contacted
CC with a specimen from a subject suspected of having a GDF-11
CC associated disorder. Antibody binding constitutes a positive result.
CC Detection is performed in muscle cells in vitro or in vivo.. The
CC antibodies may also be used in the treatment of such disorders by
CC suppressing GDF-11 activity. Antisense GDF-11 reagents may also be
CC used. Vectors are utilised in the treatment process e.g. colloidal
CC dispersion systems such as liposomes which are target specific and
CC either anatomically or mechanistically targeted.
XX
XX Sequence 126 AA;

Query Match 100.0%; Score 4; DB 17; Length 126;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMPM 4
| | | |
Db 66 ymf 69

RESULT 3
AAW23589
ID AAW23589 standard; Protein; 126 AA.
XX
AC AAW23589;

XX 10-NOV-1997 (first entry)
DE Bovine bone morphogenic protein-11.
XX
XX BMP-11; regulation; follicle stimulating hormone; FSH; contraception;
KW bone formation; cartilage formation; connective tissue formation.
XX
OS Bos taurus.

XX Key Location/Qualifiers
FT Peptide 1..17
FT Protein /label_1 Signal
FT Cleavage-site /label_2 Bone_morphogenic_protein-11
FT /note_1 "Predicted proteolytic processing sequence
FT corresponding to the consensus Arg-X-X-Arg,
FT where the signal peptide will be cleaved"

XX US5639638-A.
XX
XX 17-JUN-1997.
XX
XX 12-MAY-1993; 93US-0061464.
XX
XX 20-MAY-1994; 94US-0247907.
PR 12-MAY-1993; 93US-0061464.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Celeste AJ, Wozney JM;
PI WPI: 1997-332045/30.
DR N-PSDB; AAT74190.
XX

XX DNA encoding bone morphogenetic protein 11 polypeptide(s) - useful
PT for regulating follicle-stimulating hormone
XX
XX Claim 11; Column 25-26; 20pp; English.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2001, 20:45:17 ; Search time 25.43 Seconds
(without alignments)
3.239 Million cell updates/sec

Title: US-09-383-551b-22

Perfect score: 4

Sequence: 1 YWFM 4

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

size : 0
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	126	1	US-08-247-907A-2
2	4	100.0	126	1	US-08-452-772-2
3	4	100.0	126	2	US-08-765-875-4
4	4	100.0	126	3	US-08-795-671-4
5	4	100.0	126	5	PCT-US94-05288-2
6	4	100.0	362	1	US-08-247-907A-11
7	4	100.0	362	1	US-08-452-772-11
8	4	100.0	362	5	PCT-US94-05288-11
9	4	100.0	407	2	US-08-765-875-2
10	4	100.0	407	2	US-08-765-875-6
11	4	100.0	407	3	US-08-795-671-2
12	4	100.0	407	3	US-08-795-671-6
13	4	100.0	1311	2	US-08-540-406-4
14	4	100.0	1311	3	US-08-656-055-4
15	4	100.0	1311	4	US-08-954-668-4
16	4	100.0	1311	5	PCT-US95-13233-4
17	3	75.0	6	1	US-08-089-994A-3
18	3	75.0	6	1	US-08-429-732-31
19	3	75.0	6	1	US-08-429-732-35
20	3	75.0	6	5	PCT-US94-07605-3
21	3	75.0	7	1	US-08-429-732-21
22	3	75.0	7	1	US-08-429-732-32
23	3	75.0	7	1	US-08-429-732-33
24	3	75.0	7	1	US-08-429-732-34
25	3	75.0	7	1	US-08-429-732-36
26	3	75.0	7	1	US-08-429-732-37
27	3	75.0	7	1	US-08-429-732-38

28	3	75.0	7	4	US-09-139-802-142	Sequence 142, Appl
29	3	75.0	11	2	US-08-116-778E-8	Sequence 8, Appl
30	3	75.0	11	2	US-08-438-562-8	Sequence 8, Appl
31	3	75.0	11	2	US-08-483-528B-96	Sequence 96, Appl
32	3	75.0	12	4	US-08-819-286-25	Sequence 25, Appl
33	3	75.0	13	1	US-08-089-994A-23	Sequence 23, Appl
34	3	75.0	13	5	PCT-US94-07605-23	Sequence 23, Appl
35	3	75.0	16	1	US-07-934-656A-1	Sequence 1, Appl
36	3	75.0	16	4	US-08-819-286-26	Sequence 26, Appl
37	3	75.0	20	3	US-08-822-324-16	Sequence 16, Appl
38	3	75.0	20	3	US-08-822-324-17	Sequence 17, Appl
39	3	75.0	20	4	US-08-819-286-6	Sequence 6, Appl
40	3	75.0	20	4	US-08-819-286-27	Sequence 27, Appl
41	3	75.0	20	4	US-09-177-249-65	Sequence 65, Appl
42	3	75.0	23	1	US-08-484-635-81	Sequence 81, Appl
43	3	75.0	23	2	US-08-484-631-81	Sequence 81, Appl
44	3	75.0	23	2	US-08-827-570-81	Sequence 81, Appl
45	3	75.0	24	1	US-08-293-778-26	Sequence 26, Appl
46	3	75.0	24	4	US-08-819-286-28	Sequence 28, Appl
47	3	75.0	27	6	5190919-25	Patent No. 5190919
48	3	75.0	28	4	US-08-819-286-29	Sequence 29, Appl
49	3	75.0	31	1	US-08-393-985-21	Sequence 21, Appl
50	3	75.0	42	1	US-08-293-778-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-247-907A-2
; Sequence 2, Application US/08247907A
; Patent No. 5639638
; GENERAL INFORMATION:
; APPLICANT: WOZNEY, John
; APPLICANT: CELESTE, Anthony J.
; TITLE OF INVENTION: BMP-11 COMPOSITIONS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,907A
; FILING DATE: May 20, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: GI5205-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 876-1170
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-247-907A-2

Query Match 100.0%; Score 4; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
Db 66 YMF 69

RESULT 2
US-08-452-772-2
; Sequence 2, Application US/08452772
; Patent No. 5700911
; GENERAL INFORMATION:
; APPLICANT: WOZNEY, John
; APPLICANT: CELESTE, Anthony J.
; TITLE OF INVENTION: BMP-11 COMPOSITIONS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,772
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/247,907
; FILING DATE: 20-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LAZAR, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: G15205-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 876-1170
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-452-772-2

Query Match 100.0%; Score 4; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
Db 66 YMF 69

RESULT 3
US-08-765-875-4
; Sequence 4, Application US/08765875
; Patent No. 5914234
; GENERAL INFORMATION:
; APPLICANT: LEE, SE-JIN
; APPLICANT: MCPHERSON, ALEXANDRA C.
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US

ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,875
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,958
; FILING DATE:
; APPLICATION NUMBER: US/08/272,763
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3641
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-765-875-4

Query Match 100.0%; Score 4; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMF 4
Db 66 YMF 69

RESULT 4
US-08-795-671-4
; Sequence 4, Application US/08795671
; Patent No. 6008434
; GENERAL INFORMATION:
; APPLICANT: Se-Jin Lee and Alexandra McPherron
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,671
; FILING DATE: February 6, 1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/106001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS: